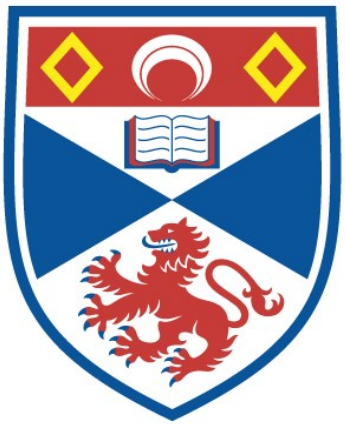
MILLIMETRE WAVE QUASI-OPTICAL SIGNAL PROCESSING AND SPREAD SPECTRUM TECHNIQUES

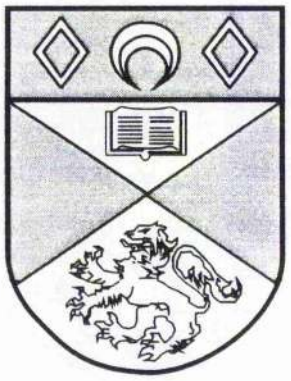
Duncan Alexander Robertson

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1995

Millimetre Wave Quasi-Optical Signal Processing and Spread Spectrum Techniques



A thesis presented by
Duncan Alexander Robertson B.Sc. (Hons)
to the
University of St. Andrews
in application for the degree of
Doctor of Philosophy

November 1994

B750

Declaration

I, Duncan Robertson, certify that this thesis has been composed by myself, that it is a record of my own work, and that it has not been accepted in partial or complete fulfilment of any other degree or professional qualification.

I was admitted to the Faculty of Science of the University of St. Andrews under Ordinance General No. 12 in October 1991 and as a candidate for the degree of Ph.D. in October 1992.

The research was carried out in the Physical Science Laboratory of St. Salvator's College, in the University of St. Andrews, under the supervision of Dr. J.C.G. Lesurf.

In submitting this thesis to the University of St Andrews I wish access to it to be subject to the following conditions:

> for a period of **2 years** from the date of submission the thesis shall be made available for use only with the consent of the head or chairman of the department in which the work was carried out.

May 1995

Duncan A. Robertson
November 1994

Certificate

I hereby certify that D.A. Robertson has fulfilled the conditions of the Resolution and Regulations appropriate to the degree of Ph.D.

Jim C.G. Lesurf
November 1994

Acknowledgements

I would like to thank Jim Lesurf for all his help, guidance and understanding throughout the course of my Ph.D. He has an amazing breadth of knowledge, some of which I hope to have gleaned over the last three years. Thanks for putting up with my sluggish ability to understand, and for all the good humour and friendship too.

Appreciation must also go towards all the members of the Millimetre Wave Group, both past and present. Mike Leeson was an invaluable source of help in the early days, and I doubt I would have got this far without the experience of my brother Malcolm, Graham Smith and Andy Harvey. Thanks too to all my friends in the Physics Department who have made life here interesting and enjoyable.

A special mention must go the Mechanical Workshop whose services really have been fundamental to this research - a big thank you to George Radley and Willie Smith in particular.

The last year and a half has been made particularly enjoyable and challenging through contact with everybody at the Passive Millimetric Imaging Group at DRA Malvern. A big "Cheers!" to Roger Appleby and Jonathan Borrill for all the work, stress and laughs!

I would also like to extend my thanks to Bal Kumar, Nigel Couch and Mike Kearney of GEC-Marconi Ltd, Hirst Research Centre, for their help and collaboration in recent months.

Thanks to Trevor Pugsley and Colin Whyte who have both done their very best to persuade me that Hi-Fi design *can* be more interesting than millimetre waves.

I am grateful for financial assistance from SERC and DRA.

Finally I should like to thank all my friends whom I have met in St. Andrews over the years and, of course, my family. You have all helped to make life here very rewarding.

Dedication

To my Parents.

Abstract

The use of quasi-optical techniques for signal processing at millimetre wave frequencies and their application to spread spectrum systems is presented in this thesis.

Millimetre waves offer a number of advantages when compared to microwave and optical signals and as a result they are finding an increasing number of applications in the area of communications.

Traditionally, millimetre wave research has been centred on scientific and experimental interests. In recent years, however, practical communications systems have been emerging at these frequencies. New technologies are being developed which aim to exploit the potential advantages of millimetric signals and the work undertaken here relates to one particular avenue - that of quasi-optics.

This thesis introduces measurement methods for assessing millimetric components and quasi-optical systems, followed by an appraisal of Gunn oscillators and their associated power supplies as suitable signal sources for communications.

A type of Gunn oscillator displaying semi-chaotic behaviour is described which may have potential as an inherently broadband source suitable for spread spectrum applications.

The application of quasi-optical signal processing methods to a prototype spread spectrum millimetre wave system is presented. The system has been shown to operate satisfactorily in a number of demonstrations. Theoretical models of the quasi-optical circuit show good agreement with experiment and an analysis of the signal-to-noise behaviour predicts the potential receiver performance. Some ideas for future work, building on that presented here, are suggested.

Table of Contents

Chapter One
Introduction to Millimetre Waves, Sources and Quasi-Optical Techniques

1.1 Millimetre Waves

Millimetre waves occupy that part of the electromagnetic spectrum which lies between 30 and 300GHz and hence they have wavelengths between 10 and 1mm. This portion of the spectrum falls between the microwave and far-infrared regions and marks a transition from waveguide techniques to the optical methods of higher frequencies.

As one moves up in frequency from microwaves to millimetre waves, waveguide propagation of signals becomes increasingly problematic. The higher frequencies lead to greater losses due to the skin-effect, and the waveguide components become harder to manufacture because of the need for very high levels of mechanical precision. Consequently, it becomes advantageous to propagate millimetre wave signals as free-space beams. However, the ray optics approach of conventional optics does not apply to millimetre waves due to the relatively long wavelengths - ray optics relies on the fact that optical elements are many hundreds or thousands of wavelengths across and this is clearly inapplicable at millimetre wavelengths.

Instead, one has to work with Gaussian beam optics. The Gaussian lenses, mirrors and other optical elements used to manipulate free space millimetre wave beams look similar to conventional optical components and so this technique is referred to as *quasi-optics*. A more detailed treatment of Gaussian beam techniques is given later in this chapter.

Whilst the work described in this thesis has been restricted to frequencies between about 70 and 150GHz, the quasi-optical techniques

employed are applicable to frequencies well into the hundreds of gigahertz.

The free space attenuation of millimetre waves is generally fairly low and is limited mainly by absorption due to water vapour. Absorption rises gradually with increasing frequency but the general trend is punctuated by a number of peaks due to molecular resonances such as the oxygen lines at 60 and 120GHz, and the water resonances at 24 and 180GHz Additionally, there are minima in the attenuation curve (called windows) at 35, 95, 140 & 220GHz which are exploited as transmission channels. Figure 1·1, below, shows the one-way attenuation characteristic with frequency.

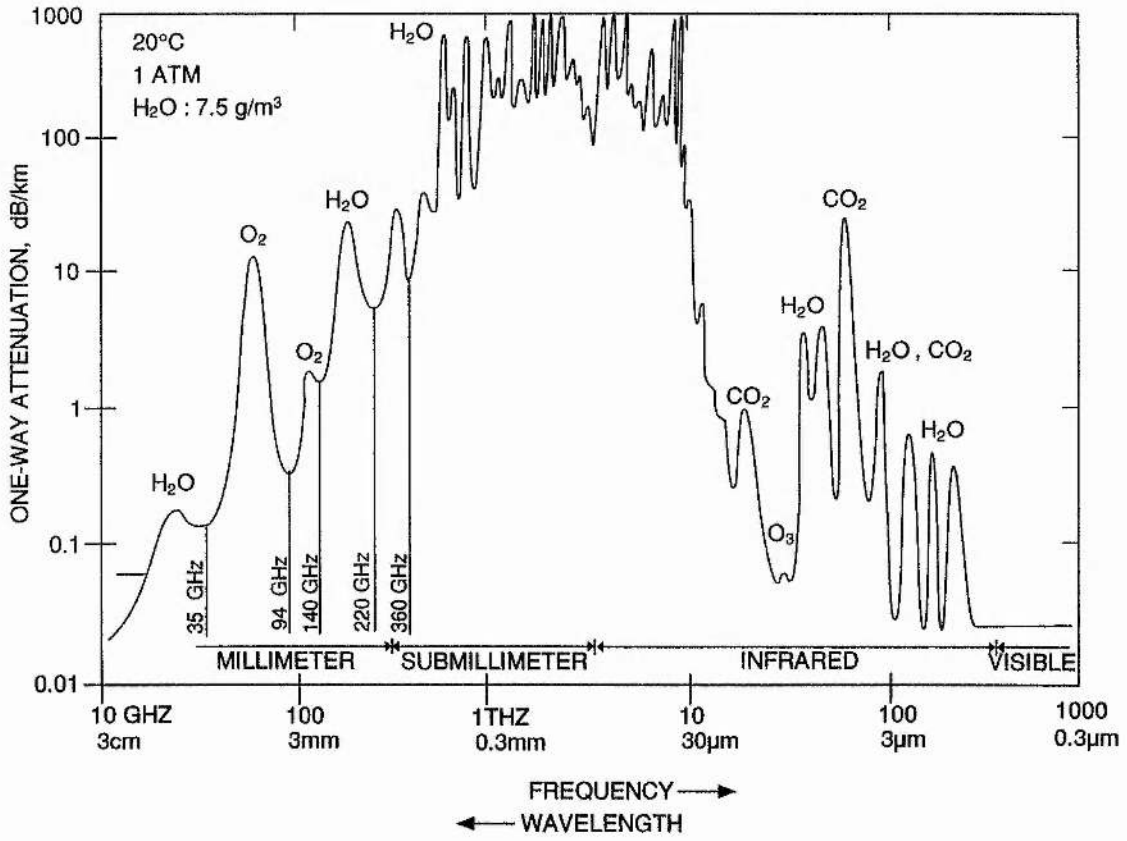


Fig. 1·1 : One-way Atmospheric Attenuation

1.2 Applications of Millimetre Waves

1.2.1 Radar and Imaging

Millimetre waves offer a number of advantages over microwaves in radar systems. The main advantages arise due to the shorter wavelengths - higher resolution, greater bandwidth, narrower beamwidth and small physical size of the hardware. Millimetre waves are also better at penetrating cloud, dust and smoke making them ideal for military radar[1,2,3,4].

The atmospheric attenuation properties of millimetre waves can also be advantageous and the windows at 35 and 95GHz are exploited for this reason.

Automotive collision avoidance radar systems operating at millimetric frequencies are being developed for several major car manufacturers. The European Community has allocated the 76-77GHz band for automotive radar but demonstration systems have been made at a number of frequencies between 60 and 94GHz. Ultimately, collision avoidance radar will be incorporated into intelligent cruise control systems which would maintain a vehicle's position relative to the road boundaries and relative to other vehicles [5,6,7,8,9,10].

Imaging techniques can be either passive, which relies on the natural thermal emissions of an object, or active which uses an illuminating beam with the image being produced by detecting the reflected power. Both systems are based on radiometer receivers in linear or two-dimensional arrays. As with all forms of radiometry, there is a trade-off between resolution and refresh rate. High resolution images are produced at 94GHz but the integration time per pixel limits the frame rate to a few seconds[11,12]. Alternatively, video rates are possible but at lower resolution[10].

Millimetre wave sensors are able to penetrate fog, smoke and

cloud better than infrared systems and they also offer increased angular resolution over microwave systems.

1.2.2 Communications

The potentially high bandwidths and compact physical size of millimetre wave systems mean that they are ideal as carrier frequencies for communication channels. Millitech offer a range of millimetre wave radios for voice, data and video links operating at frequencies from 24 to 320GHz[10]. Such systems are ideal for use in areas already crowded with lower frequency signals and are compact and portable.

The oxygen resonance at 60GHz is exploited to provide short range secure communications channels[13]. The high atmospheric absorption gives rise to lower sidelobe levels yielding improved local isolation, less mutual interference and a lower probability of intercept. Hence, air-to-air communications between fighter aircraft operating in localised groups may use 60GHz links.

Certain satellite communications operate in the millimetre wave region due to the low probability of intercept and the broad transmission bandwidths available[14].

The rapidly developing field of spread spectrum communications, currently being exploited at lower frequencies for digital radio broadcasts[15,16] and mobile telephones[17,18], will inevitably move to higher frequencies in order to fulfil the requirement of very wide transmission bandwidths. Additionally, as digital systems move into gigabit-per-second data rates, millimetre wave technology will play an increasing role in many aspects of communications. An introduction to spread spectrum techniques appears in Chapter 6.

1.2.3 Plasma Diagnostics[19]

The electron-cyclotron resonant frequency, and its harmonics, of a magnetically confined fusion plasma lie in the range 50 to

200GHz. The electron density profile can be determined with millimetre wave reflectometers[20] and heterodyne radiometery measurements of electron-cyclotron emission yields the electron temperature[21]. Thomson scattering measurements are made to study ion and alpha-particle distribution within the plasma[22].

Additionally, the electron-cyclotron resonance can be exploited for rf heating of the plasma. This requires very high power millimetre wave sources and has led to the development of megawatt gyrotrons[23].

1.2.4 Electron Spin Resonance

Electron spin resonance (ESR) spectrometers measure the absorption of continuous, single-frequency radiation by a sample whilst sweeping an applied magnetic field through resonance. The resonances are a function of electron structure in the sample and ESR systems are used to probe the electronic structure of atoms. The first ESR systems operated at microwave frequencies but the use of millimetre wave sources gives enhanced spectral resolution and yields more information on the hyperfine electronic structure of the sample under test[24,25]. High frequency ESR spectroscopy is of particular use for studying short-lived free radicals. It is also possible to map the electron spin density in a sample and this technique is known as ESR imaging[26].

1.2.5 Remote Sensing

The strong interaction of millimetre waves with atmospheric gases and particulates means that the propagated wave serves as a diagnostic tool for sensing meteorological properties of the atmosphere. Typical parameters which are studied at millimetric frequencies include temperature, water vapour content, wind velocity, cloud composition and rain distribution[27].

Recently, a great deal of effort has been directed to

monitoring levels of stratospheric ozone and so-called "greenhouse gases". Satellite and airborne receivers are used to detect halogens and oxides of nitrogen and hydrogen which are thought to be responsible for the destruction of the ozone layer[28,29].

1.2.6 Materials Measurement

Dispersive Fourier Transform Spectroscopy (DFTS) is a widely used technique which allows the determination of the optical properties of solids, liquids and gases at millimetre and submillimetre wavelengths. DFTS uses broadband interferometry to measure the complex transmission or reflection spectra of a specimen. The sample is placed in one arm of an interferometer and the attenuation and phase shift due to the sample allows computation of its refractive index and complex permittivity. Such measurements have been used to determine the optical properties of a vast range of substances[30].

1.2.7 Radio Astronomy

Molecular clouds in interstellar space can be surveyed with millimetre and submillimetre wave instruments. The composition of interstellar gas and dust clouds can be determined by using heterodyne receivers which examine the rotational spectra of molecules and free-radicals.

The high frequencies and sensitivities required for radio astronomy have been a driving force in the development of higher frequency oscillators, multipliers and mixers with improved low-noise performance. Much of the technology developed in this way will find its way into many of the other applications already described[31,32].

1.3 Sources of millimetre wave radiation

Whilst many types of sources exist, the work presented here concentrates on solid-state sources, and emphasis will be placed, in particular, on Gunn-effect oscillators.

1.3.1 Lasers and Vacuum Devices

Optically pumped molecular lasers (e.g. $C^{13}H_3F$, CH_3I) and free-electron lasers (FELs) can be attractive as very high power mm wave sources (up to 1MW pulsed). They typically cover the range from 250GHz to 4.5THz but tend to operate essentially as single frequency sources and have limited tuning propeties[33].

Vacuum devices generate rf power due to the dynamic interaction of electrons with a magnetic field. Broadly speaking, a beam of electrons is injected into a resonant cavity, which is surrounded by a static magnetic field, and accelerated towards an anode. The resulting gyromagnetic interaction causes a transfer of kinetic energy from the electrons to the resonant rf field and coherent radiation is produced. Backward-wave oscillators (BWOs) and travelling-wave tubes (TWTs) utilise a periodic metallic structure ("slow-wave" structure) which interacts with the electron beam to generate an rf electromagnetic field.

Klystrons operate up to about 200GHz and can deliver hundreds of milliwatts of power. Magnetrons work at lower frequencies but offer power outputs of around one kilowatt. BWOs are wideband tunable devices and are available for use throughout the mm and submm ranges. They produce tens to hundreds of milliwatts[34]. Inspired by the need for megawatts of power for plasma heating, the gyrotron has achieved the highest output powers of any millimetre wave source[35].

All the sources mentioned above, whilst endowed with impressive output powers, tend to have limited lifetimes and require

large high-voltage power supplies. The devices themselves are also bulky and expensive and suffer from poor noise performance. For compact, reliable sources which offer good low-noise performance one has to look to semiconductor devices.

1.3.2 Solid-State Devices

There are three principal types of solid-state millimetre wave source, namely the IMPATT diode, the Gunn diode and the quantum-well (QW) device. IMPATT oscillators operate up to about 200GHz at powers of several milliwatts, but can reach hundreds of milliwatts at lower frequencies or in pulsed operation[36]. Gunn devices only work up to 150GHz and typically produce a few tens of milliwatts. Both IMPATTs and Gunn diodes are long-established and are used in many commercial and military applications. A more novel and as yet experimental device which may prove to be of particular use at higher frequencies is the quantum-well device (or resonant tunnelling diode - RTD). Such devices have been shown to oscillate at up to 712GHz, which is the highest ever reported oscillation frequency for a solid-state device[37]. The principal disadvantage of RTDs is their tiny power output, typically only a few microwatts.

Solid-state sources offer a number of advantages over lasers and vacuum tube oscillators, in particular longevity, low power consumption, low-noise operation and compactness.

A more detailed look at the operation and behaviour of Gunn diodes is presented below.

1.3.2.1 The Gunn Effect

Gunn diodes take their name from J.B. Gunn who first reported field-dependent current oscillations in a wafer of GaAs[38]. They are bulk-effect semiconductor devices which exhibit a region of negative differential resistance (ndr) above a threshold value of

applied electric field. The mechanism creating the ndr region is a field-induced transfer of conduction band electrons from a high-mobility central valley to a low-mobility satellite valley. It is for this reason that such devices are generally known as transferred-electron devices (TEDs).

The two materials which are principally employed as TEDs are n-type GaAs and InP. As the applied field is increased, electrons scatter from a high-mobility, low energy state in the central Γ valley to a low-mobility, high energy state in an upper *L* satellite valley, as shown in Fig. 1·2.

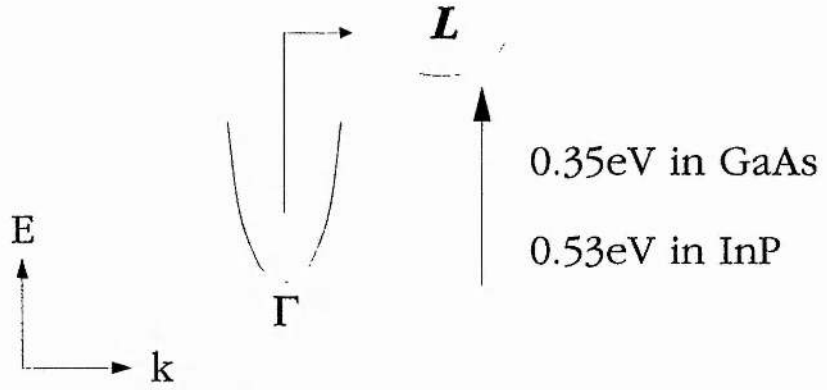


Fig. 1·2 : Schematic of Band Structure in a TED

Since the electron drift velocity is related to their mobility, this effect represents a field-dependent velocity distribution in the material. As the applied field is proportional to the applied voltage, V, and the electron velocity is proportional to the current through the device, I, this leads to the familiar nonlinear I-V curve of the Gunn diode. Fig.1·3 shows the region of negative differential resistance.

If the device is placed in a suitable circuit and biassed into the ndr region, it will oscillate thereby converting dc energy into an rf output. The efficiency with which this energy conversion takes place is typically about 2% for InP and less than 1% for GaAs, the remainder being lost as heat.

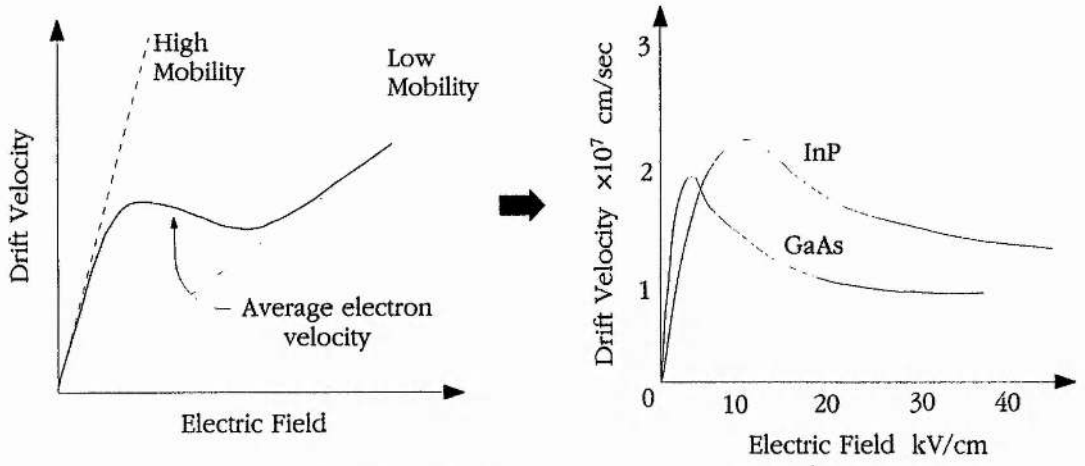


Fig. 1·3 : Electron drift velocity versus electric field

Gunn diodes operate in a number of different modes, the simplest being the classical domain mode which describes low-frequency operation. As electrons are injected into the material from the cathode, they slow down as they scatter into the satellite valley and subsequent electrons catch up, creating an accumulation zone. Electrons in front of this region are accelerated towards the anode, forming a depletion region. This dipole domain travels along the device at the drift velocity and appears at the anode as a current pulse. Hence, this is a transit time effect and the oscillation frequency is determined by the device length.

At higher frequencies the diodes are too short to operate as transit time devices and they operate in an accumulation mode or in a complex hybrid mode. The upper limit of operation for GaAs diodes is about 75GHz whilst it is around 140GHz for InP. Above 75GHz, GaAs diodes usually operate in second harmonic mode, in which the inherent nonlinearity of the diode causes harmonic generation[39].

1.3.2.2 Gunn Oscillators

At frequencies above 10GHz, the parasitic effects associated with a Gunn diode and its packaging become very important and it is necessary to use a well defined resonant circuit to match the

impedance of the device. The impedance of a Gunn diode is typically a few ohms and this has to be matched to a waveguide output with a characteristic impedance of a few hundred ohms (at 94GHz, the impedance of WG27 waveguide is 484Ω).

The Gunn oscillators employed at St. Andrews are all of the resonant cap and post, coaxial cavity design derived by Smith[40], based on work by Carlström *et al*[41] - Fig. 1·4.

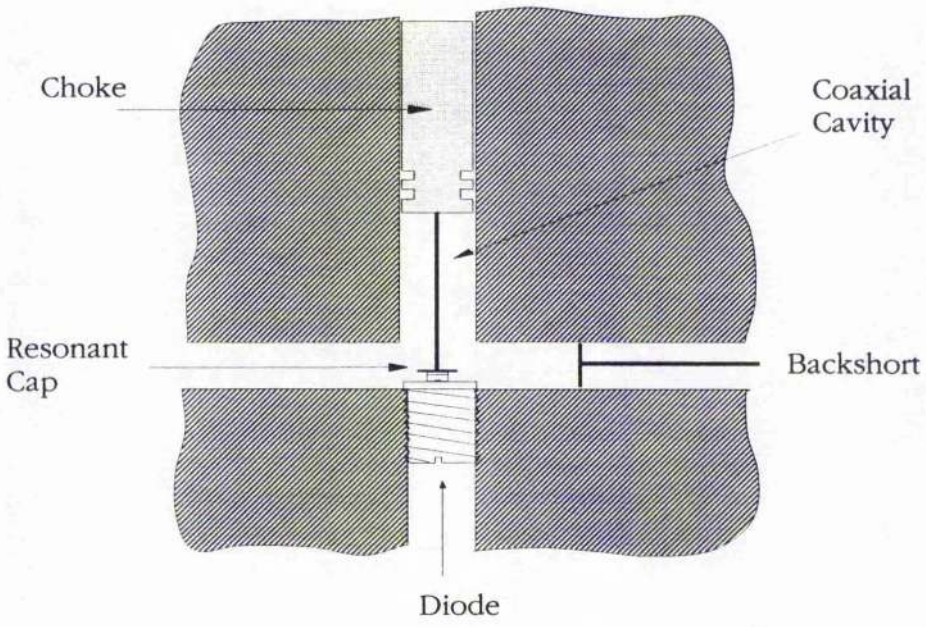


Fig. 1·4 : Coaxial cavity with resonant cap

The diode is mounted flush with the broad wall of the waveguide and sits below a cylindrical cavity. Most of the oscillators used here operate in second harmonic mode such that the cut-off frequency of the waveguide is above the fundamental frequency of the diode and only the second harmonic (and occasionally the third harmonic) can travel down the waveguide. The fundamental resonates in a quasi-coaxial mode entirely within the cavity and hence very high Q-factors (at the fundamental) are achieved to help match the impedance of the diode. The second harmonic is only weakly coupled into the cavity.

The resonant cap, which sits on top of the diode package,

acts as an impedance transformer (or radial-line transformer) at the second harmonic frequency.

The frequency of operation is tuned by altering the height of the coaxial cavity. This is achieved by adjusting the height of the choke, which is spring-loaded against a micrometer screw. A second spring inside the choke ensures that as this adjustment is made, the cap and post are kept in good contact with the diode, for the electrical connection to the bias supply is made via this route. The backshort controls the termination of the second harmonic and thus controls the output power coupling.

The cap dimensions are critical to the operation of the oscillator. A smaller cap diameter increases the oscillation frequency and the cap thickness is usually kept small (typically less than 0·2mm) to maximise power output. The exact cap dimensions are determined empirically for each oscillator.

Detailed descriptions of the design and performance of Gunn oscillator blocks of this type can be found in[40,41,42,43].

1.3.2.3 Electronic Tuning of Gunn Oscillators

The frequency of oscillation of a Gunn diode, for any given cavity setting, can be altered by adjusting the applied bias voltage. Changing the bias voltage not only changes the applied field but also the power which the device has to dissipate as heat and hence alters the temperature of the active layer. Gunn diodes have strong frequency dependence and power conversion efficiency dependence with operating temperature. Henceforth, the frequency and power are tuned simultaneously with bias.

Since the current-voltage and frequency-temperature characteristics are nonlinear, so too is the frequency-bias tuning curve. The power also varies nonlinearly with bias. This frequency tunability makes Gunn devices attractive as voltage-controlled oscillators (VCOs) but the nonlinearities mean they are far from

perfect as modulated sources.

At low modulation frequencies, the frequency tuning is principally due to thermal effects induced by the varying amount of power which is injected into the device. Frequency tuning is also determined by the capacitances associated with the diode. These electronic effects, masked by the thermal properties at low modulation rates, become significant as the rate of change of applied bias increases.

A transferred electron device can be represented, albeit in a greatly simplified form, by the following equivalent circuit, Fig. 1·5.

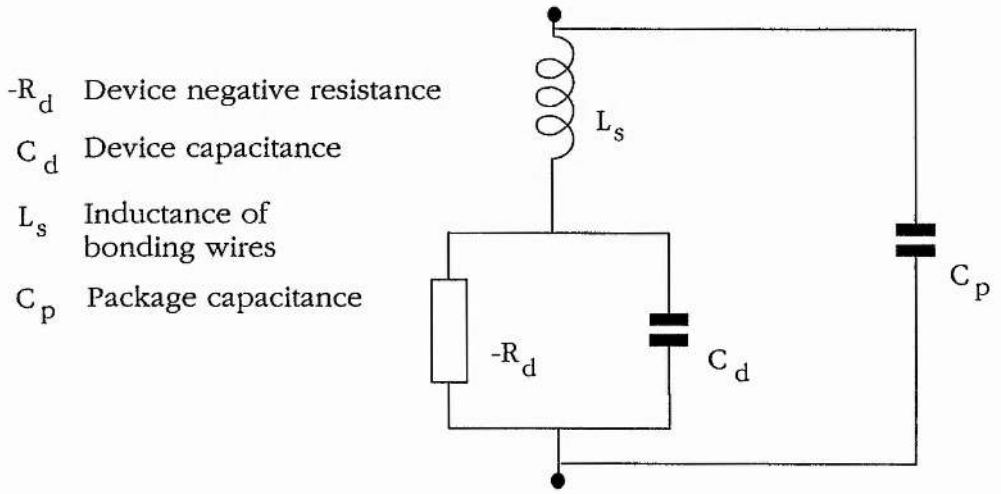


Fig. 1·5 : Equivalent circuit of a TED

The device capacitance, C_d, is a voltage dependent capacitance which, in conjunction with the parasitics L_s and C_p, alters the rf impedance of the circuit and consequently changes the oscillation frequency.

A deliberate enhancement of this property is made in a varactor-tuned oscillator. In such a device, a voltage dependent capacitor, or varactor, is mounted near the Gunn diode and biassed from a separate supply. The capacitance of the varactor is now the circuit element which can be used to tune the frequency of

oscillation. The advantage of this type of tuning is that the varactor draws very little current so that the modulation bandwidth is much higher. It is also possible to achieve tuning ranges of several GHz whilst conventional bias tuning is typically limited to sweeps of only a few hundred MHz[44,45].

1.4 Introduction to Gaussian Beams

As was briefly mentioned at the beginning of this chapter, it is not practical to treat millimetre wave free-space beams with a conventional ray optics approach. Ray optics requires that the optical elements are many hundreds or even thousands of wavelength across, and this would lead to impractically large millimetre wave optics. Therefore, to be able to manipulate millimetric signals with compact free-space beams one has to adopt the approach of Gaussian Beam Mode (GBM) theory.

GBM theory makes two physical assumptions about the optical system being described:

- the radiation is moving in a paraxial beam (which essentially travels along an axis but with some diffraction), the cross-sectional size of which is too small to be treated as plane parallel.
- the radiation can be represented as a scalar field.

This then allows the wave equation for the signal to be expressed in its simplified scalar form

$$\nabla^2\psi + k^2\psi = 0 \qquad 1.1$$

where $k = 2\pi f / c$ and ψ represents the scalar field appropriate for the beam. c is the speed of light and f is the signal frequency.

The solution to this equation represents a set of orthogonal spatial modes described by Gauss-Hermite polynomials (when using the cartesian coordinates). It has become a standard result, derived in texts

on GBM theory[46,47] and it will simply be quoted here. The fundamental mode is the most commonly used and further discussion of free-space beams will be limited to this case. The E-field for the fundamental mode is

$$\psi_x = E_o\sqrt{\frac{2}{\pi\omega^2}}\exp\left(\frac{-r^2}{\omega^2}\right)\exp\left(\frac{-jkr^2}{2R}\right)\exp(-j\Phi_o)\exp(j\{kz - 2\pi ft\}) \quad 1.2$$

with a similar expression for ψ_y, and $r^2 = x^2 + y^2$.

The term inside the square root normalizes the field pattern to have power $|E_o|^2$ when integrated over a transverse plane. The first exponential shows that the amplitude variation across the beam is Gaussian in profile with the radial half-width of the beam, called the beamwidth ω. The beamwidth is defined as the distance from the beam axis at which the field amplitude falls to $1/e$ of its axial value. The second exponent describes the spherical nature of the phase fronts, which have radii of curvature R, and the third describes the phase slippage Φ_o, relative to a plane wave at the same frequency. The last term is the time varying, travelling electromagnetic wave of frequency f.

The beamwidth along the axis of propagation has a minimum known as the beamwaist, ω_o. Conventionally, the axis of propagation is z and the coordinate origin is placed such that $z = 0$ at the plane of the beamwaist. The variation of ω, R and Φ_o with z are found from the derivation of eqn. 1.2 and are given by

$$\omega^2 = \omega_o^2\left\{1 + \left[\frac{z\lambda}{\pi\omega_o^2}\right]^2\right\} \quad 1.3$$

$$R = z\left\{1 + \left[\frac{\pi\omega_o^2}{z\lambda}\right]^2\right\} \quad 1.4$$

$$\Phi_o = \arctan\left\{\frac{z\lambda}{\pi\omega_o^2}\right\} \quad 1.5$$

The transition from near- to far-field occurs when the radius of curvature reaches its minimum value and is known as the Rayleigh Range (or confocal distance), given by

$$z_R = \frac{\pi\omega_o^2}{\lambda} \qquad 1.6$$

Fig. 1·6 depicts how a Gaussian beam propagates along z, showing the radial Gaussian amplitude dependence and the variation of ω and R along z.

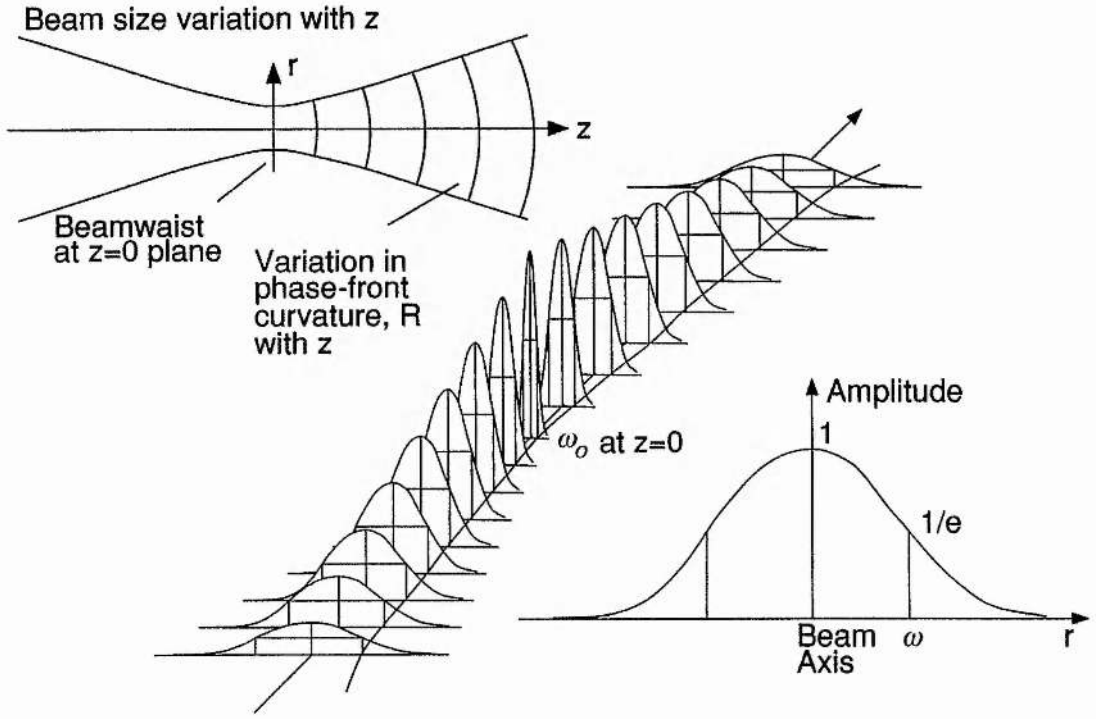


Fig. 1·6 : Gaussian Beam Propagation

1.5 Fundamentals of Quasi-Optics

In order that signal processing can be performed with free space millimetre wave beams they must be manipulated with quasi-optical components. The design of optical components and circuit hardware has been treated in detail elsewhere so I will present just some of the most important points.

Fundamental mode beams are best launched into free space with corrugated feed horns[48]. These feeds are very efficient and nearly 98% of the output power is contained in the fundamental mode.

Since the radial amplitude dependence of a such a beam is Gaussian it theoretically has infinite extent. Clearly, there has to be a practical limit placed on the size of optical apertures and this will inevitably cause some truncation. All else being equal, the width of a beam increases with increasing wavelength, so for any given circuit, the size of the optical aperture will be governed by the lowest operating frequency. The amount of truncation can be limited and is often negligible. In practice, it is customary to set the minimum aperture diameter equal to three times the largest beamwidth which will pass through the plane of the aperture, *i.e.*

$$d_{min} = 3\omega_{max} \qquad 1.7$$

This results in an output beam which has all but 2% of the input power contained in the fundamental mode. This is considered to be an acceptable level of error due to truncation

As indicated in Fig. 1·6, Gaussian beams diverge with distance and the rate of divergence increases with decreasing beamwaist (eqn. 1.3). Therefore, a limit is approached beyond which the beam cannot pass through an aperture without suffering appreciable truncation. Henceforth, the beam has to be confined within the apertures of the optical elements through which it passes. This is accomplished by focussing the beam down to keep it to a manageable size and this can be performed by using either off-axis mirrors or dielectric lenses[47].

Dielectric lenses alter the phase front curvature of a beam in such a way that for a given input beamwaist at a given distance from the lens, an output beam is generated which has a different beamwaist located at a different distance from the lens - Fig. 1·7.

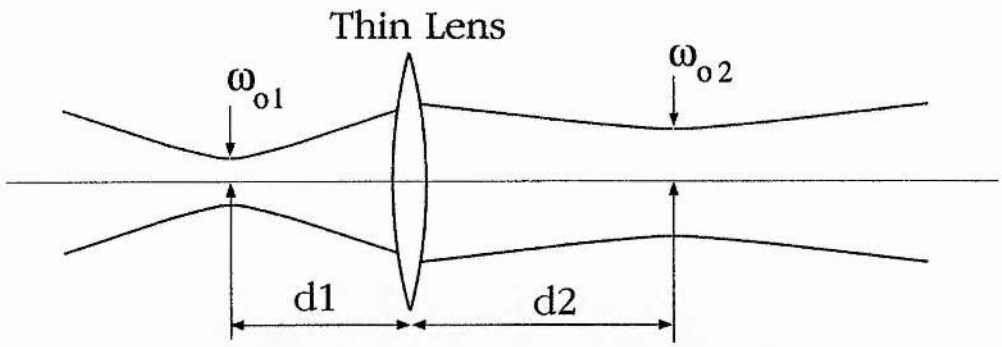


Fig. 1·7 : Gaussian Beam focussing by a thin lens.

Lenses suffer from absorption in the material and from reflection at the air/dielectric interface. The former problem can only be minimised by using thin lenses and low-loss material, whilst the latter is addressed through the use of anti-reflection matching layers. These could be quarter wavelength layers of additional material of the appropriate refractive index, or as is more usual, a blazed lens surface. The blazing can take the form of concentric grooves or a pattern of narrow holes which are machined out of the lens surface to create a pseudo-layer of material with a different refractive index. The lenses made at St. Andrews are all made from high density polyethylene (HDPE) and are blazed[49]. Blazing does, however, introduce a frequency dependency and hence the technique is relatively narrowband.

Off-axis mirrors have potentially very low ohmic losses, but are physically harder to incorporate into a circuit design. They will not be considered in this thesis. Lenses preserve the linear nature of a signal path through a circuit since they are in-line elements. Refocussing a beam repeatedly through a series of optical elements gives rise to the notion of a "beam waveguide" - Fig. 1·8.

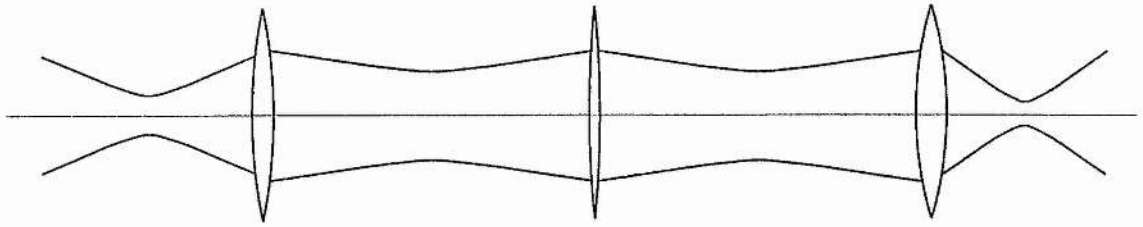


Fig. 1·8 : Beam Waveguide.

Rewriting eqn. 1.3 as a quadratic in ω_o^2 gives us

$$\omega_o^2 = \frac{\omega^2}{2} \pm \frac{\sqrt{\omega^4 - \left[\frac{2z\lambda}{\pi}\right]^2}}{2} \qquad 1.8$$

which shows that for a given set of λ, z and ω it is possible to form two different sized beamwaists with a single lens. Since the square root term must be real we have

$$z \leqslant \frac{\pi\omega^2}{2\lambda} = z_{max} \qquad 1.9$$

where z is the distance between the beamwaist and the position at which the beamwidth is ω. At $z = z_{max}$ the two solutions are coincident and this is known as the maximum throw. It is not possible to form a beamwaist at values of z greater than the maximum throw. For $z < z_{max}$ the larger of the two waists is called the parallel waist and the smaller is known as the focussed waist[47].

The condition of maximum throw has a major bearing on the design of quasi-optical circuits and the size of the optical components, in particular. In a complicated circuit the distance between lenses must not exceed maximum throw if truncation is to be avoided. As shown above, it is the parameters λ, z and ω which determine the maximum throw. Therefore, when circuit hardware is being designed one has to take account of the following factors:-

- the minimum signal frequency
- the optical path length through the circuit
- the spacing between the lenses
- the total number of lenses (and hence absorption losses)

- aperture size

For laboratory applications, total system size is rarely an issue and optimum performance is more important. Consequently, the general purpose optical components used in St. Andrews are relatively bulky. They are based around aluminium half-cubes which fit onto a gridded baseplate - a design which allows maximum flexibility for constructing different circuits whilst providing repeatably accurate alignment of components[49,50]. For specific applications it is possible to reduce significantly the size of quasi-optical circuits[51]. The issue of circuit performance versus system size is addressed in Chapter 9.

REFERENCES

[1] Skolnik. M.I., *Introduction to Radar Systems*, 2nd Ed., McGraw-Hill, 1980, pp. 560-566.
[2] Currie, N.C., & Brown, C.E., *Principles and Applications of Millmeter-Wave Radar*, Artech House, 1987.
[3] Johnston, S.L., "MM-Wave Radar Challenges and Benefits EW Applications", *Microwave Syst. News & Commun. Tech.*, June 1986, pp. 95-110.
[4] Johnston, S.L. "MM-Wave Radar: The New ECM/ECCM Frontier", *Microwave Jnl.*, May 1984, pp. 265-271.
[5] Stove, A.G., "80GHz radar for cars", *Proc. IMechE FISITA 1992 - Safety, The Vehicle and the Road*, pp. 115-121.
[6] Stove, A.G., "Obstacle detection radar for cars", *Electronics & Commun. Eng. Jnl.*, Oct. 1991, pp. 232-240.
[7] Binns, P., "FMCW radar: a low cost sensor for automotive applications", *Proc. 1991 Eur. Microwave Conf.*, pp. 147-152.
[8] Mallinson, P., & Stove, A.G., "Car avoidance radar at 94GHz", *Proc. 1989 IMechE Automotive Electronics Conf.*, pp. 297-302.
[9] 60GHz car radar with 300m range, *Hughes Millimeter-Wave Products Catalog*, 1991, p. 5.
[10] Browne, J., "MM waves aid commercial applications", *Microwaves & RF*, July 1992, pp. 113-116.
[11] Appleby, R., Gleed, D.G., Anderton, R.N., Lettington, A.H., "High performance passive millimetre wave imaging", DRA Malvern, Worcs WR14 3PS.
[12] Wilson, W.J., Howard, R.J., Ibbott, A.C., Parks, G.S., Ricketts, W.B., "Millimeter-Wave Imaging Sensor", *IEEE Trans. Microwave Theory & Tech.*, **MTT-34** (10), Oct. 1986, pp. 1026-1035.
[13] 60GHz secure radio system, *Alpha Industries Millimeter Wave Components & Subsytems Catalog*, 1989, p. 8-5.
[14] 44GHz uplink & 20GHz downlink mobile, tactical sat-comms, *Hughes Millimeter-Wave Products Catalog*, 1991, p. 5.
[15] Fox, B., "When Broadcasting goes Digital", *Hi-Fi News & Record Review*, Oct. 1991, pp. 31-37.
[16] Reinhart, E.E., "Telecommunications Special Report", *IEEE Spectrum*, Feb. 1992, pp. 24-26.
[17] ibid., pp. 27-29.
[18] Schilling, D.L., Pickholtz, R.L., & Milstein, L.B., "Spread spectrum goes commercial", *IEEE Spectrum*, Aug. 1990, pp. 40-45.
[19] Luhmann, Jr., N.C., "Instrumentation and Techniques for Plasma Diagnostics: An Overview", *Infrared & MM Waves*, **2** (Ch. 1), Ed. K.

Button, Academic Press, 1979, pp. 1-65.

[20] Costley, A.E., "Development of reflectometry for plasma density measurements at JET", *Proc. 18th Intl. Conf. IR & MM Waves*, Colchester, Sept. 1993.

[21] Salmon, N.A., "First electron temperature edge measurements on the ASDEX Upgrade tokamak using a heterodyne radiometer", *Proc. 18th Intl. Conf. IR & MM Waves*, Colchester, Sept. 1993.

[22] Fessey, J.A., Hoekzema, J.A., Hughes, T.P., "The receiver system of the fast ion and alpha particle diagnostic at JET", *Proc. 18th Intl. Conf. IR & MM Waves*, Colchester, Sept. 1993.

[23] Kreisher, K.E., & Temkin, R.J., "High-Frequency Gyrotrons and Their Application to Tokamak Plasma Heating", *Infrared & MM Waves*, **7** (Ch. 8), Ed. K. Button, Academic Press, 1983, pp. 377-485.

[24] Lebedev, Y.S., *Modern Pulsed and C-W ESR*, Ed. L. Kevan & M.K. Bowman, Wiley, 1990, p. 365.

[25] Smith, G.M., Lesurf, J.C.G., & Riedi, P., "A Millimetre-Wave Quasi-Optical Electron Spin Resonance Spectrometer", *Proc. 19th Intl. Conf. IR & MM Waves*, Sendai, Japan, Oct. 1994.

[26] Eaton, S.S, & Eaton, G.R., *Modern Pulsed and C-W ESR*, Ed. L. Kevan & M.K. Bowman, Wiley, 1990, p. 405.

[27] Alsthuler, E.E., "Millimeter-Wave Propagation and Remote Sensing of the Atmosphere", *Infrared & MM Waves*, **9** (Ch. 4), Ed. K. Button, Academic Press, 1983, pp. 177-239.

[28] Kerridge, B.J., "Potential of millimetre/sub-millimetre heterodyne instruments to sound atmospheric composition", *Proc. 18th Intl. Conf. IR & MM Waves*, Colchester, Sept. 1993.

[29] Nett, H., Crewell, S., & Künzi, K., "A submillimeter-wave sensor for trace gas studies in the middle atmosphere", *Proc. 18th Intl. Conf. IR & MM Waves*, Colchester, Sept. 1993.

[30] Birch, J.R. & Parker, T.J., "Dispersive Fourier Transform Spectrometry", *Infrared & MM Waves*, **2** (Ch. 3), Ed. K. Button, Academic Press, 1979, pp. 137-271.

[31] Newport, R.W., "The UK/NL millimetre wave telescope", *Endeavour*, **8** (4), 1984, pp. 159-165.

[32] Richards, P.L. & Greenberg, L.T., "Infrared Detectors for Low-Background Astronomy: Incoherent and Coherent Devices from One Micrometer to One Millimeter", *Infrared & MM Waves*, **6** (Ch. 3), Ed. K. Button, Academic Press, 1982, pp. 149-207.

[33] Detemple, T.A., "Pulsed Optically Pumped Far Infrared Lasers", *Infrared & MM Waves*, **1** (Ch. 3), Ed. K. Button, Academic Press, 1979, pp. 129-184.

[34] Kantorowicz, G. & Palluel, P., "Backward Wave Oscillators", *Infrared*

& MM Waves, **1** (Ch. 4), Ed. K. Button, Academic Press, 1979, pp. 185-212.

[35] Dialetis, D. & Chu, K.R., "Mode Competition and Stability Analysis of the Gyrotron Oscillator", *Infrared & MM Waves*, **7** (Ch. 10), Ed. K. Button, Academic Press, 1983, pp. 537-581.

[36] Midford, T.A. & Bernick, R.L., "Millimeter-Wave CW IMPATT Diodes and Oscillators", *IEEE Trans. Microwave Theory & Tech.*, **MTT-27** (5), May 1979, pp. 483-492.

[37] Brown, E.R., Söderström, J.R., Parker, C.D., Mahoney, L.J., Molvar, K.M. & McGill, J.C., "Oscillations up to 712GHz in InAs/AlSb resonant-tunneling diodes", *Appl. Phys. Lett.*, **58**, 1991, pp. 2291-2293.

[38] Gunn, J.B., "Microwave Oscillations of Current in III-V Semiconductors", *Solid State Commun.*, **1** (4), 1963, pp. 88-91.

[39] Eddison, I.G., "Indium Phosphide and Gallium Arsenide Transferred-Electron Devices", *Infrared & MM Waves*, **11** (Ch. 1), Ed. K. Button, Academic Press, 1984, pp. 1-59.

[40] Smith, G.M., *TEOs at MM-Wave Frequencies and their Characterisation using Quasi-Optical Techniques*, Ph.D. Thesis, St. Andrews, 1990.

[41] Carlström, J.E., Plambeck, R.L. & Thornton, D.D., "A Continuously Tunable 65-115GHz Gunn Oscillator", *IEEE Trans. Microwave Theory & Tech.*, **MTT-33** (7), July 1985, pp. 610-619.

[42] Haydl, W.H., "Fundamental and Harmonic Operation of Millimeter-Wave Gunn Diodes", *IEEE Trans. Microwave Theory & Tech.*, **MTT-31** (11), Nov. 1983, pp. 879-889.

[43] Robertson, M.R., *Novel Solid-State Millimetre-Wave Sources*, Ph.D. Thesis, St. Andrews, 1992.

[44] Sultan, N., "Theoretical Analysis & Experimental Verification of Wideband Tunable Radial Line Millimeter Waveguide Voltage Controlled Oscillators", *Proc. 15th Intl. Conf. Infrared & MM Waves*, 1990, pp.94-97.

[45] Cohen, L.D., "A Millimeter-Wave Third-Harmonic, Gunn VCO with Ultrawideband Tuning", *Proc. IEEE MTT-S*, 1991, pp.937-938.

[46] Goldsmith, P.F., "Quasi-Optical Techniques", *Infrared & MM Waves*, **6** (Ch. 5), Ed. K. Button, Academic Press, 1982, pp. 277-343.

[47] Lesurf, J.C.G., *Millimetre-wave Optics, Devices & Systems*, Adam Hilger / IOP, 1990.

[48] Wylde, R.J., "Millimetre-wave Gaussian beam-mode optics and corrugated feed horns", *Proc. IEE, H*, **131** (4), Aug. 1984, pp. 258-262.

[49] Harvey, A.R., *A Quasi-Optical Millimetre Wave Compex Impedance Bridge*, Ph.D. Thesis, St. Andrews, 1990.

[50] Leeson, M.J., *The Application of Quasi-Optical Techniques to Millimetre Wave Radar*, Ph.D. Thesis, St. Andrews, 1992.

[51] Kumar, B., GEC-Marconi Ltd., Borehamwood, Private Communication.

Chapter Two
General millimetre wave measurement techniques

Much of the work described in this thesis is based around a number of standard measurement techniques and in this chapter I will present an introduction to these methods and the equipment employed.

2.1 Oscillator Characterisation

All Gunn oscillators fabricated at St. Andrews are characterised for their frequency and power output over their full tuning range. Quite often, the first characterisation shows that the oscillator does not achieve the required performance and it has to be modified, either by changing the diode or by altering the dimensions of the cap. This process may have to be repeated a number of times until the necessary performance is reached.

Oscillators are characterised by measuring their frequency and power for a range of choke positions (cavity heights). For any given choke position, the backshort is tuned to give maximum power. Data are recorded separately for tuning up and down, as the small amount of mechanical backlash in the choke assembly gives rise to a hysteresis in the tuning behaviour. The bias voltage is kept constant for these measurements. Bias tuning behaviour is dealt with in Chapter 4.

Power is measured with one of two power meters - either an Anritsu ML83A or a Boonton 4220. Frequency can be measured in two ways. The simplest technique uses an EIP 578 Source Locking Microwave Counter which gives direct readings of frequency in W-band. Alternatively, the frequency can be determined with a Martin-Puplett Interferometer which is equipped with an electronic fringe counter that gives a reading of signal wavelength[1]. See Fig. 2·1.

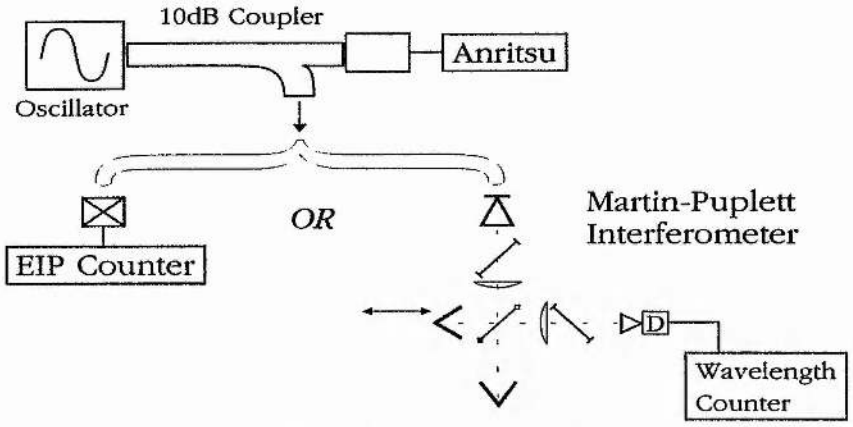


Fig. 2·1 : Oscillator Characterisation Apparatus

A number of oscillators were characterised in the manner described and whilst the detailed results are not directly relevant here, there was agreement with the behavioural trends described elsewhere[1,2]. Frequency jumps, power dips and tuning hysteresis were frequently observed. The oscillation frequency increased with decreasing cap diameter, but at reduced power output.

A comparison was made of the two frequency measurement techniques for the same oscillator. It was found that the interferometer gave a smoother tuning characteristic than that obtained with the EIP, particularly at high frequencies. If the signal is of very low power, or contains significant harmonics, the EIP can jump to the "wrong" harmonic of its local oscillator and thereby give an incorrect reading. Using the interferometer, one can visually monitor the detector signal for distortion, on an oscilloscope, and the spectrum can be computed to display signal harmonics. Using the EIP there is no way of checking the purity of the oscillator signal, but the readings can be made in considerably less time.

Fig. 2·2 shows an example of a tuning characteristic obtained for a GEC diode DB637/6/1#1 in a full-height WG27 block. The post and cap diameters were 1.0 and 2.20mm respectively. Bias was 5.0V.

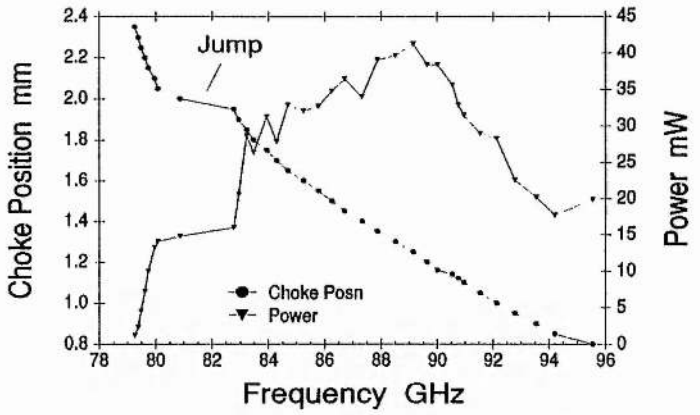


Fig. 2·2 : Example Oscillator Tuning Characteristic

2.2 Antenna Pattern Measurement

A facility for measuring the far-field antenna pattern of a feedhorn was available due to the work of Webb[3]. This system consists of a receiver mounted on a long aluminium arm which traces a circular arc in the horizontal plane, through the far-field of the feedhorn under test. The arm is driven by a precision dc servo motor and gearbox. A potentiometer mounted on the pivot provides a reference voltage indicating the position of the arm. The source is chopped and a lock-in amplifier used to recover the signal before sampling and processing by computer.

The receiver on the arm consists of a pyroelectric detector developed by Webb[4] and a smooth walled conical feedhorn. The NEP of the detector is about 30nW $Hz^{-1/2}$ at a 20Hz modulation rate.

Mike Leeson used this system to evaluate novel tunable Potter horns and simple conical feeds[5] but the small off-axis power readings computed from the analogue-to-digital convertor (ADC) values were found to fluctuate substantially thus obscuring detail in the antenna sidelobes. Sidelobe power levels are typically 20dB down on the central lobe so the detail at this level was being lost in the quantisation noise of the ADC. Additionally, it was found that the lock-in amplifier

was adding noise to the signal (probably due to inadequate power supply regulation and smoothing).

In response to this problem, I constructed a logarithmic amplifier to convert the detected signal into decibels so that the sampled readings would already be in units of power.

2.2.1 Log Amp Design

The circuit is based around an SSM-2100P log/antilog amplifier IC from PMI. This device can be configured with the addition of a few external components to take either the log or the antilog of the ratio of two input voltages. Either of the inputs can also be set to a fixed reference voltage. The output is also scalable by choosing particular resistor values. The circuit design incorporates buffered inputs and output, switchable reference voltages, and switchable output scaling. Fig. 2·3 shows the circuit.

The SSM-2100P was configured as a logarithmic amplifier according to the data sheet[6] such that linearity is achieved by accurately trimming two input resistors $R_{in} = R_{ref}$, so that operation follows the expression

$$V_{out} = K\log_{10}\left(\frac{V_{ref}}{V_{in}} \cdot \frac{R_{in}}{R_{ref}}\right) = K\log_{10}\left(\frac{V_{ref}}{V_{in}}\right) \qquad 2.1$$

where K is the output scale factor.

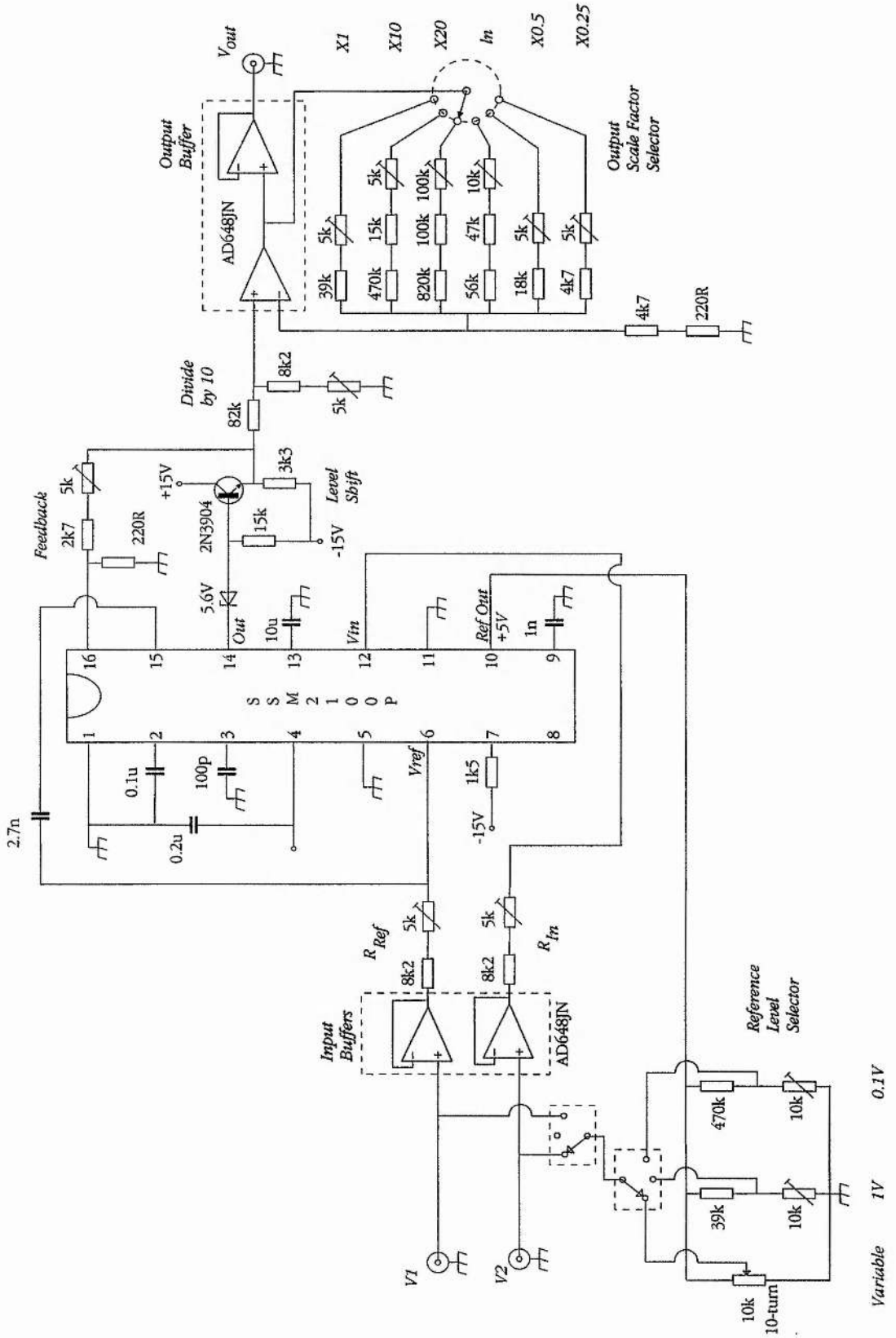


Fig. 2·3 : Log Amplifier Circuit

Either V_{in} or V_{ref} can be switched to one of three reference voltages derived from pin 10 of the SSM-2100P (1V, 0.1V or a variable voltage in the range 0-5V) whilst the other is connected to the input socket, or both voltages can be derived from external signals (log ratio operation).

Since the output of the IC is positive only, a level shifting circuit is incorporated to give bipolar output which exploits the full 39dB dynamic range of the ADC[7]. The level shift and feedback sections are derived from the manufacturers data sheet. A 1:10 potential divider is included to permit overall scale factors of less than unity. The non-inverting output amplifier (gain always greater than one) has switched feedback to give the required scale factors of ×1, ×10, ×20, natural logs (ln), ×0.5 and ×0.25.

The buffer ICs are Analog Devices AD648JN (dual op-amp) chosen for their low input bias current.

2.2.2 Log Amp Results

The circuit was tested over three decades of dc test voltages for each combination of input and reference voltage, and output scale factor. The results were all accurate in following a logarithmic relationship. Any small deviations may be due to the untrimmable dc offset of the AD648JNs. An excellent example of the response of the circuit is shown in Fig. 2·4 overleaf.

The SSM-2100P is a dc log amp and a brief test with an ac input confirmed that logarithmic behaviour only occurred at signal frequencies which were so slow as to be quasi-static. This apparent limitation turned out to be an advantage in the context of the antenna pattern measurement system. If the receiver arm is swept at a low rate, the slow response of the log amp provides a degree of signal averaging. Additionally, the power supply noise from the lock-in amplifier is well suppressed.

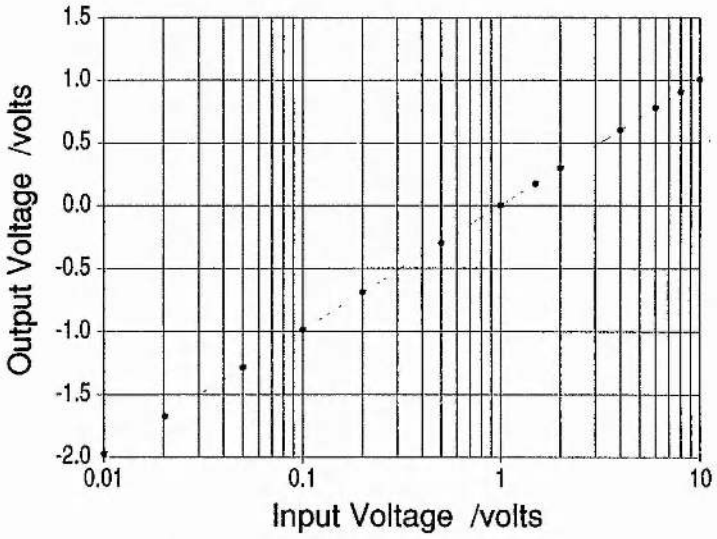


Fig. 2·4 : Log Amp Characteristic (V2 referenced to 1.00V)

In addition to these advantages, the log amp increased the dynamic range of the antenna measurements to approximately 50dB and greatly reduced the problems due to quantisation errors. An example of an antenna pattern taken with the system incorporating the log amp is shown in Fig. 2·5.

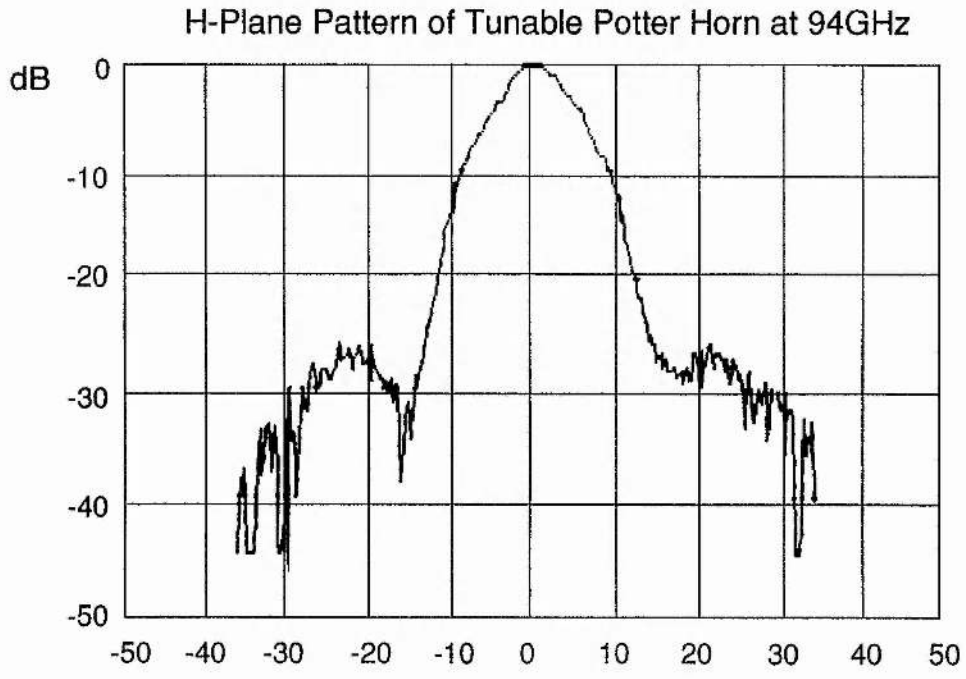


Fig. 2·5 : Example Antenna Pattern

2.3 Free-Space Power Meter Calibration

It is possible to measure the power of a free-space millimetre wave beam using a Terahertz Absolute Power Meter from Thomas Keating Engineering[8]. The unit is suitable for free space beams up to 30mm in diameter with frequencies in the range 30GHz to 3THz. It has an NEP of $5\mu WHz^{-1/2}$.

The sensor head is a photoacoustic detector which consists of a closed air-filled cell formed by two closely spaced parallel windows. A thin metal film lies in the middle of the gap which absorbs very close to 50% of the incident power, the remainder being transmitted or reflected. The beam must be 100% amplitude modulated in the range 10-50Hz and the fluctuating power level causes changes in temperature and hence pressure in the cell. This modulation of the pressure is detected by a sensitive microphone. The microphone signal is recovered with a lock-in amplifier referenced to the source chopper.

The modulated pressure change is closely proportional to the total absorbed power and depends negligibly on the power distribution over the film. Calibration is achieved by ohmic heating of the film which has a resistance in the range 110-200Ω. Ideally, the calibration is a nulling technique incorporating a lock-in amplifier which ensures that the heating voltage is applied exactly out of phase with the chopped rf signal. The rf power can then be deduced from the amount of ohmic power supplied to the film which gives a null output signal. Whilst this system can be supplied by the manufacturer it was not available in this case and a simpler substitute had to be devised.

2.3.1 Power Meter Control Circuit

The power meter controller was to be based on sinusoidal heating of the film. If the rf power is chopped at f_o then the applied heater voltage should be at $\frac{1}{2}f_o$ since $P=V^2/R$. Thus any electrical pick-up between the heater and the microphone will appear at $\frac{1}{2}f_o$

and the desired signal at f_o can be selected with a bandpass filter.

The bandpass filtering was to be done by half of an MF10 switched capacitor filter IC[9], clocked at 64× the centre frequency f_o. The heater voltage could also be derived from f_{clk} by dividing it by 128 and low-pass filtering (with the other half of the MF10) to select the fundamental signal at $\frac{1}{2}f_o$. A signal at f_o for electronic chopping of the source could also be easily obtained from the clock.

The circuit design also incorporated a switched gain (0-50dB) preamplifier for the microphone signal and a stepped attenuator (0-40dB) for the heater voltage. Additionally, a chopper monitor signal and phase shifter were required.

Unfortunately, the circuit design was beset by numerous problems and I had insufficient time to remedy them. In particular, the low-pass filtering of the f_{clk}÷128 signal was insufficient and the heater signal was far from sinusoidal. In retrospect, my design of the control circuit had tried to incorporate too many features and as a result, the only usable parts of the design were the preamp and the chopper monitor.

Since the power meter is so sensitive, measuring low power levels is made very difficult because of acoustical noise and vibration being picked up by the microphone. 50Hz mains pick-up is also a problem on the detector signal.

2.3.2 Comparison with waveguide power meter

The free-space power meter was compared with our Anritsu ML83A unit for a range of power levels. The principal difficulty encountered was that of how to chop the source as required by the Thomas Keating (TK) meter whilst measuring continuous power with the Anritsu. If the chopper was placed too close to the TK meter the draught cause by the rotating blades appeared as significant acoustical noise. Measurements below 100μW were almost impossible due to noise.

However, I obtained reasonable results for moderate power levels. The Anritsu was mounted on the -10dB arm of a waveguide coupler and a corrugated feedhorn on the other. The chopper and TK meter were placed close to the feedhorn to ensure all of the beam passed through the sensor window. The power measured by the TK unit was calculated from the sinusoidal ohmic heating required to give the same signal level as generated by the rf power. The power is given by

$$P = 2 \times \frac{(V_{pk})^2}{R_h} = 4\frac{V_{rms}^2}{R_h} \qquad 2.2$$

where the factor of 2 is present because only half the power is absorbed in the film, of resistance R_h. V_{rms} is the heater voltage.

A graph of the comparison of the powers measured with the two meters is shown in Fig. 2·6.

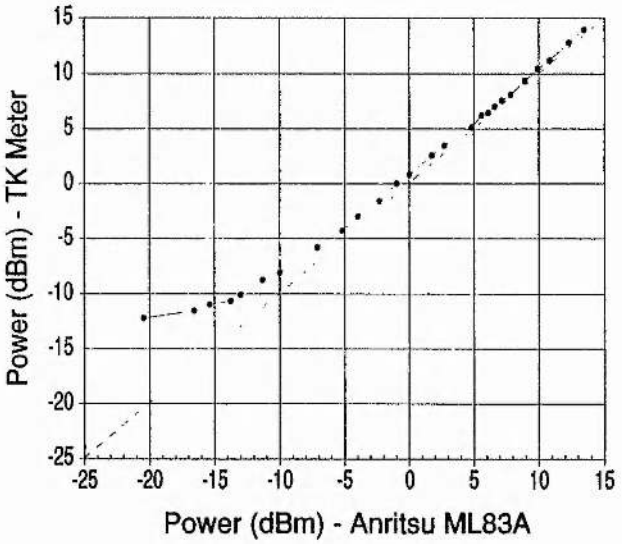


Fig. 2·6 : Comparison of free-space and waveguide power meters

Note the reasonable linearity above a few milliwatts and the increasing discrepancy below about 100μW. This deviation may be due to a combination of two effects. Firstly, as has been noted above, the TK meter suffers from noise problems below these power levels and secondly, the Anritsu being on the -10dB arm of the

coupler was actually measuring power levels a factor of ten smaller. However, I am not certain as to the exact cause of this effect. The overall behaviour does show that the free-space meter consistently reads higher than the waveguide sensor.

2.3.3 Calibration by ohmic heating

For quick power measurements the detector output of the free-space meter was calibrated for a range of known heater powers. In this way, I constructed look-up tables or graphs of incident power versus detector voltage. Applied heater power was in the range 100mW to 200μW and the detector voltage was measured with both an rms meter and by reading the peak-to-peak voltage off an oscilloscope. The modulation rate was 20Hz.

The results were very consistent and accurate to a few percent. The responsivity of the meter is specified only to ±30%. The calibration yielded the following rules of thumb for determining the incident power which generates a given detector voltage (either rms or peak-to-peak when modulated at 20Hz):-

$$Power\ in\ mW = 43 \times V_{rms} \quad or \quad 15 \times V_{pp} \qquad 2.3$$

Power levels below 200μW can be interpolated using these figures but with increased uncertainty due to noise.

2.4 Martin-Puplett Interferometer

One of the most widely used instruments in this work is the polarising Michelson, or Martin-Puplett Interferometer[10,11]. Principally, it is used to measure the frequency spectrum of an oscillator under test and its implementation for this use is described below.

2.4.1 System Overview

Whilst a Martin-Puplett Interferometer (MPI) was already

available in St. Andrews[1], it was a physically bulky system and the computer control (by a BBC B micro) was slow and unsophisticated. I therefore developed a new MPI from the existing instrument. This new system uses the quasi-optical building blocks designed by Harvey[7] and the control is performed by an Acorn Archimedes computer. Whilst the optics were immediately available, interface electronics and control software had to be implemented.

The interferometer is required to drive one roof mirror up and down on a linear track whilst an encoder generates pulses indicating the position of the mirror. The detector signal is sampled at particular intervals of mirror position and an interferogram is recorded. Taking a Fast Fourier Transform (FFT) of the interferogram yields the signal spectrum.

2.4.2 Interface Electronics

The Archimedes computer was equipped with a Wild Vision ADC-1208 card[12] which consists of an 8-bit digital I/O port and an eight input 12-bit ADC which can sample at up to 167kSa/sec. The control of the mirror and handling of the encoder signals is dealt with by the digital I/O port whilst the ADC is used to sample the detector signal.

The mirror moves on a lead screw driven by a precision dc motor whose speed is controlled by the applied dc voltage. The direction of travel is determined by the polarity of the motor voltage. The mirror drive voltage is derived from a flip-flop which can be preset or cleared either by computer signals or by microswitches at the ends of the mirror travel. The motor drive also incorporates a speed control, an on/off switch and a computer controlled on/off relay.

The linear encoder[13] mounted on the mirror produces two orthogonal pulse trains to monitor distance, and a single pulse signal to indicate the middle of travel. Each pulse train has a period of

8μm and they are XOR'd together to give a signal with a period of 4μm. This is then divided by eight and the new 32μm-period signal determines the sampling interval. This corresponds to 21.25 pulses per millimetre of travel. The 4μm-period signal is also divided by five and passed to a fringe counting wavelength meter[1].

The middle position indicator is a single 8μm pulse which could be easily missed by the computer so it is used to toggle a flip-flop which can then be read and reset by the computer.

In the encoder circuit diagram, shown in Fig. 2·7 below, the PB*n* lines are the digital controls of the I/O port.

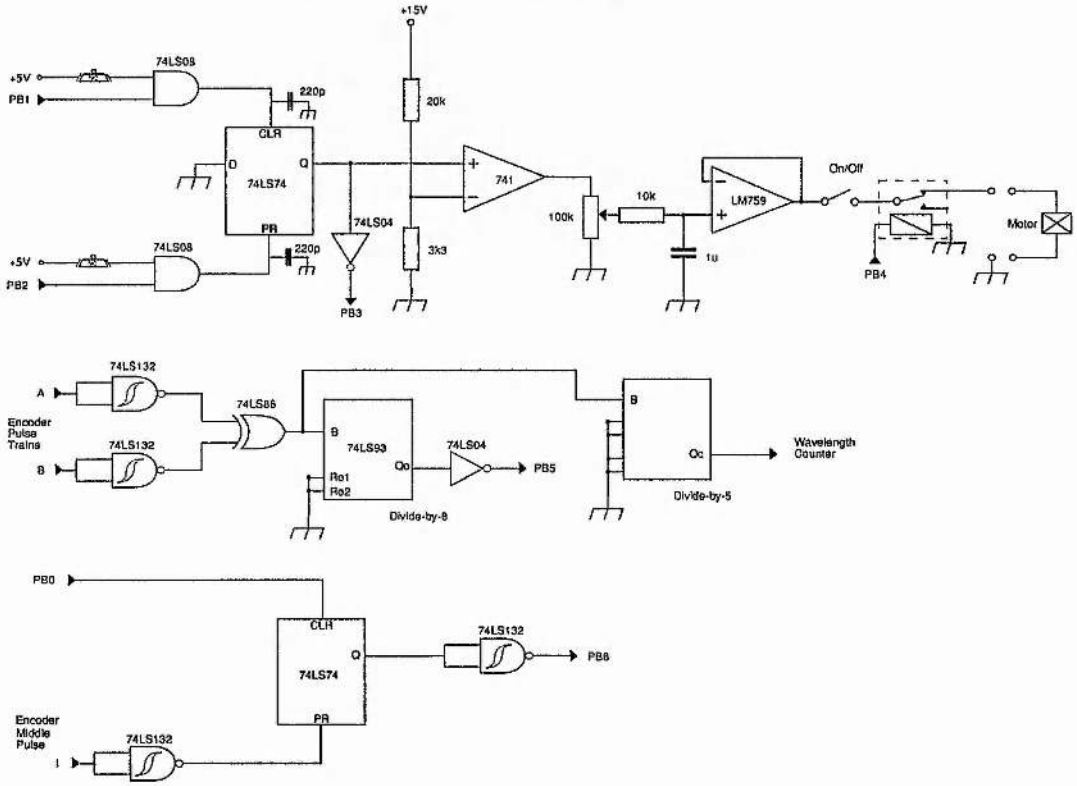


Fig. 2·7 : Interferometer Interface Circuit

2.4.3 Software

The control software was written by Martin Dunstan and contains the following principal features.

The mirror is driven up and down over a predetermined

length of its travel, centred around the middle point and the detector voltage is sampled every 32μm. Each scan consists of 4096 pulses of data with 312 buffer pulses at each end. Scans in opposite directions are stored in separate arrays and scans can be overlaid by adding successive data sets. The FFT of the data for a chosen scan is computed and the power spectrum and phase plot are displayed. This can then be saved to disk. The resolution of the FFT is 1.1435GHz which is determined by the length of the scan. Unfortunately, the mirror travel limits the resolution to this value.

2.4.4 Results

The phase repeatability of the MPI was checked over nine scans by recording the phase of the spectral peak of a free-running Gunn oscillator. The frequency of the source was stable to within the resolution of the FFT.

If two waves which add together are $A\cos\beta$ and $A\cos(\beta+\phi)$ where ϕ is the rms phase error between them, then if $\phi=0$, the waves add up exactly to give $2A\cos\beta$ which is a maximum value of 2A, when $\beta=0$. Now if $\phi\neq0$ we have

$$A\cos\beta + A\cos(\beta+\phi) = A(1+\cos\phi) \quad ; \text{for } \beta = 0 \qquad 2.4$$

So as a fraction of maximum we have

$$Phase\,Error = \frac{(1+\cos\phi)}{2}$$

$$\rightarrow 10.\log_{10}\left[\left(\frac{1+\cos\phi}{2}\right)^2\right] dB \qquad 2.5$$

In this way the phase deviation was calculated as -0.003dB for scans in and -0.001dB for scans out. This is an acceptably good result.

An example of a signal spectrum obtained with the new interferometer is shown in Fig. 2·8.

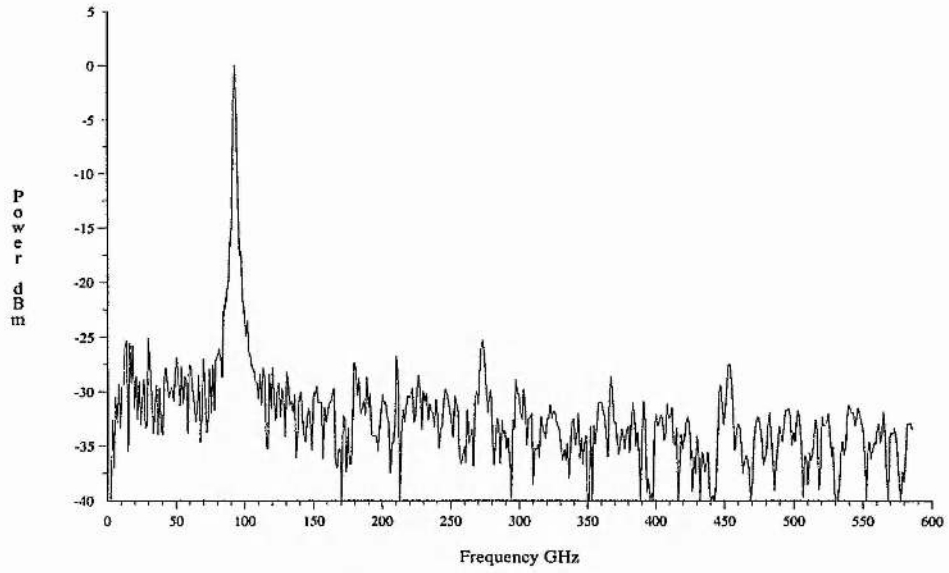


Fig. 2·8 : Oscillator Spectrum

2.5 Schottky-diode Detector Characterisation

A number of Schottky-diode detectors were fabricated at St. Andrews and I characterised their frequency response and tuning behaviour. The diodes used were pill-packaged devices from Farran Technology Ltd. of Cork[14] and the detector blocks were direct derivatives of the St. Andrews Gunn oscillator blocks described by Smith[1]. Hence, the frequency sensitivity of such a detector is controlled by altering the height of the coaxial cavity above the diode package. Power coupling is tuned with a sliding waveguide backshort. Bias is brought to the device in the same way as for the Gunn oscillators - down the centre of the cavity by a metal post which is spring contacted onto the lid of the pill-pack. The bias of ~ 0.84V @ 70μA is supplied from a low-noise battery supply[5]. In all the measurements I describe here, the output voltage of the detector includes a preamplifier of gain 6.5. Thus with zero rf power incident, the output measures 5.47V and this decreases with increasing rf power. So, the following graphs appear to be upside-down. The diodes are not designed to be operated without bias.

The measurement set-up consisted of a Gunn oscillator coupled through a quasi-optical isolator[3] to a 50:50 wire grid splitter. Half the power was collected by an Anritsu ML83A power meter as a reference and the other half was incident on the detector under test. The frequency of the source was measured with an EIP 578 microwave counter.

2.5.1 Response with Backshort Tuning

The source power and frequency were kept constant as the detector backshort position was altered. Measurements for detector block "A" with Farran diode SD013-014-01#3 are shown in Fig. 2·9. The source frequency was 90.16GHz and the detector choke was set to 2.82mm.

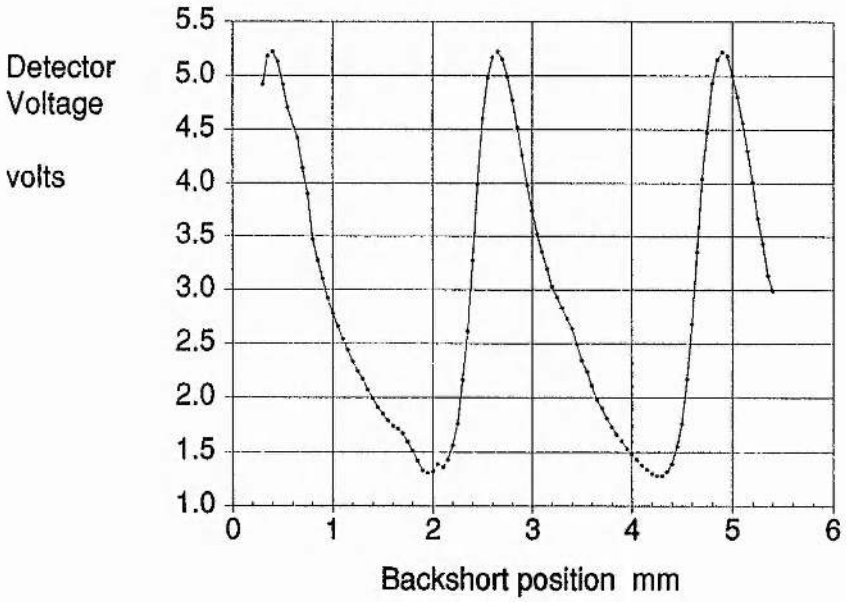


Fig. 2·9 : Detector Backshort Response

The graph clearly shows the cyclical nature of output voltage with backshort position and the non-symmetrical tuning which is common in oscillator characterisation[1]. Fig. 2·10 below shows both detector output and source power as monitored with an Anritsu power meter. Results are for the same block and diode as in Fig. 2·9. Source freq. = 91.318GHz; Detector Choke = 2.80mm

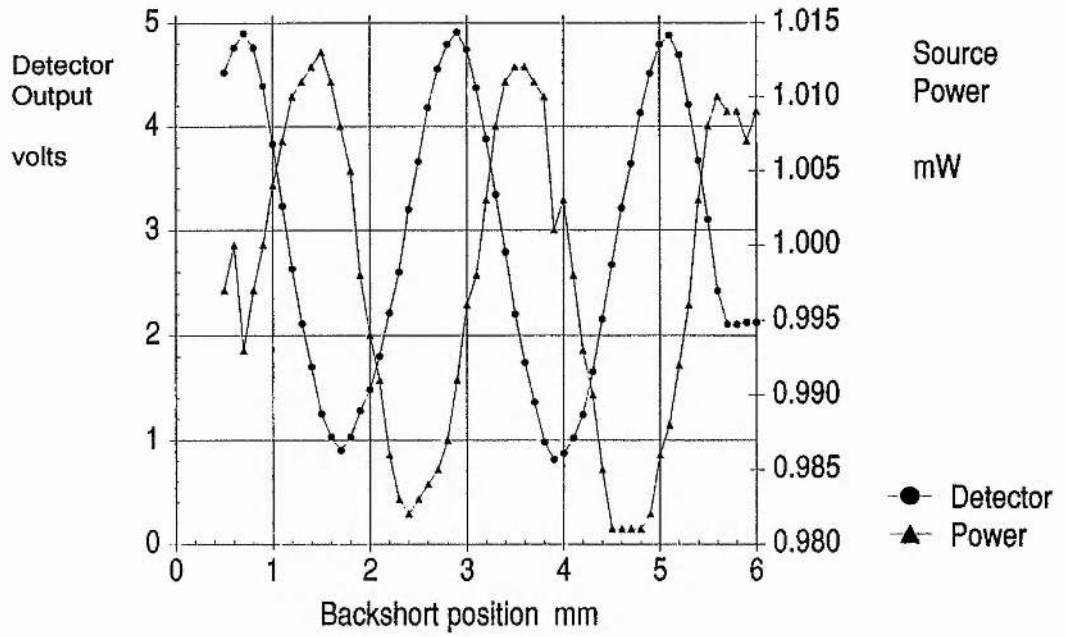


Fig. 2·10 : Detector response and source power variation

Note how the source power varies cyclically with detector output. This is due to the VSWR of the detector changing as the backshort is moved, causing a varying power level to be reflected back into the optics and being monitored by the Anritsu. However, the fluctuation in power is only of the order of three percent and is thus not very significant.

2.5.2 Response with Choke Position

The source power and frequency and the detector backshort were all kept fixed as the detector choke postion was altered. Fig. 2·11 below shows the asymmetric 'peak' in the response curve, and the corresponding variation in source power (about 10%), due again to changing VSWR. Results are for the same block and diode as in Fig. 2·9. Source freq. = 87.557GHz; backshort position = 0.82mm. The calculated peak sensitivity is 740mV/mW, at a choke position of 2.75mm.

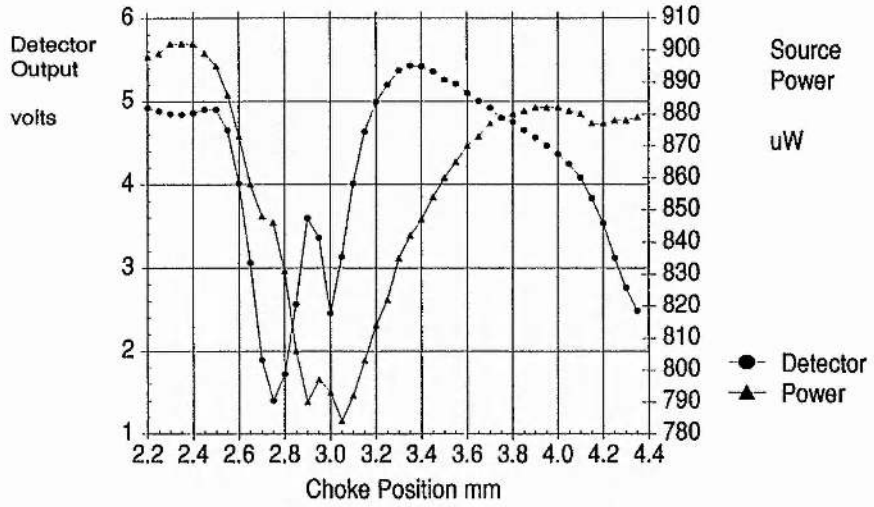


Fig. 2·11 : Detector Choke Response

2.5.3 Swept Frequency Response

The detector choke and backshort were left unaltered as the oscillator was tuned over a 10GHz frequency range. The power was adjusted at each frequency increment to be as constant as possible and the variation was no more than two percent. Fig. 2·12 shows the frequency response of the same detector as measured above. Choke = 2.75mm; Backshort = 3.27mm; Source power = 515µW. The peak sensitivity is 905mV/mW at 87.75GHz.

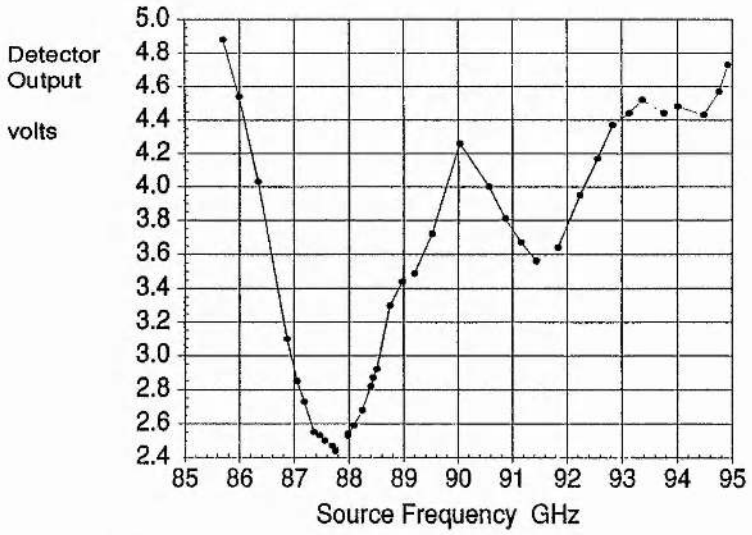


Fig. 2·12 : Frequency Response over 10GHz

The asymmetry of the frequency response of the detector is clear and the hump around 90GHz is similar to that shown in Fig. 2·11. In order to study the region of the main peak more closely, I altered the frequency of the oscillator by bias tuning and normalised the power level by using a rotatable wire-grid polariser. In this way, I achieved frequency sweeps of about 700MHz around the peak with power fluctuations of less than one percent. The detected power level was 653μW. Fig. 2·13 shows the frequency response obtained in this way for the same block and diode as in Fig. 2·9. The peak sensitivity is 590mV/mW at 85.53GHz. Note the bump around 85.45GHz - this effect was seen with other block/diode combinations but its cause is not understood. It might be due to standing wave effects in the measurement optics.

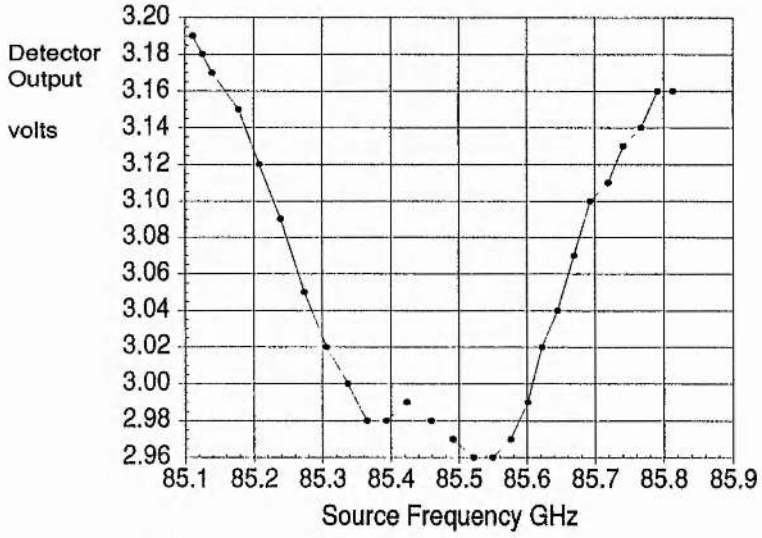


Fig. 2·13 : Frequency response around peak

The measurements made on these Schottky-diode detectors indicate that they are very frequency selective and have asymmetric frequency response curves. Comparison of a number of diode/block combinations showed a great variation in response due mainly to variations in the diodes. Sensitivities could vary by a factor of two just between diodes of the same type and tuning behaviour was not consistent. General trends in asymmetry and variable VSWR were

evident and the tuning characteristics of the blocks agreed with those of the oscillators of the same type.

The detectors are not suitable for broadband work and are certainly not useful as true power (square-law) detectors. However, for single frequency use they are adequately sensitive for many applications and have performed reasonably well as mixers.

2.6 Novel Detector Block

The ready contacted pill-pack diodes referred to above are now no longer in production but the uncontacted diode chips are still available[14]. These semiconductor chips contain many Schottky diodes in a mesa array and a whisker contact is required to complete the circuit. This arrangement requires a new design of waveguide detector block so a prototype was produced, based on the Sharpless wafer[15].

The diode chip is mounted on a one millimetre diameter brass pin mounted flush with the lower broad wall of the waveguide. A similar brass pin with a phosphor bronze whisker is mounted in the opposite, top wall of the guide and contact with the diode is achieved across the waveguide gap.

A brass wafer was designed to mount the two pins across a half-height WG27 waveguide and to provide electrical contact from the whisker pin to an SMA connector on the top edge. The waveguide was chosen to be half-height to try to match the waveguide and diode impedances. The lower edge of the wafer is cut away to allow adjustment of the diode pin. Additionally, this reduces the length of the 1mm hole which has to be drilled. It is not practical to try to drill a 1mm diameter hole longer than about 10mm.

The wafer is then located in a slot machined in an aluminium block which contains a half-height to full-height waveguide taper for the output, and a half-height waveguide with sliding contacting backshort for adjustable power coupling. The wafer is accurately held

in the slot by means of two precision located dowels which run from one side of the block, through the wafer to the other side of the block.

Both the wafer and the block are constructed in upper and lower halves to facilitate machining of the waveguide sections. Each pair of halves is dowelled and bolted together.

I am greatly indebted to Willie Smith of the Physics Department Mechanical Workshop for contructing the detector and for his useful advice in the design stage.

Unfortunately, the new detector remains untested due to insufficient time and the lack of certain whisker contacting apparatus. A diagram of the detector appears in Fig. 2·14 and the whole detector assembly appears in the photograph overleaf.

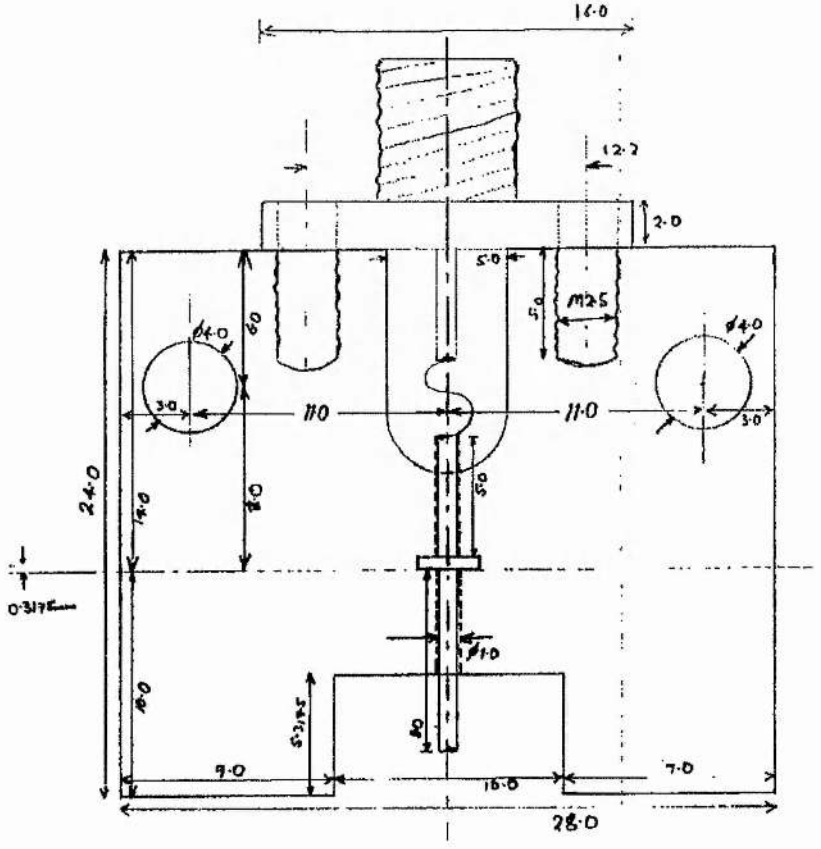
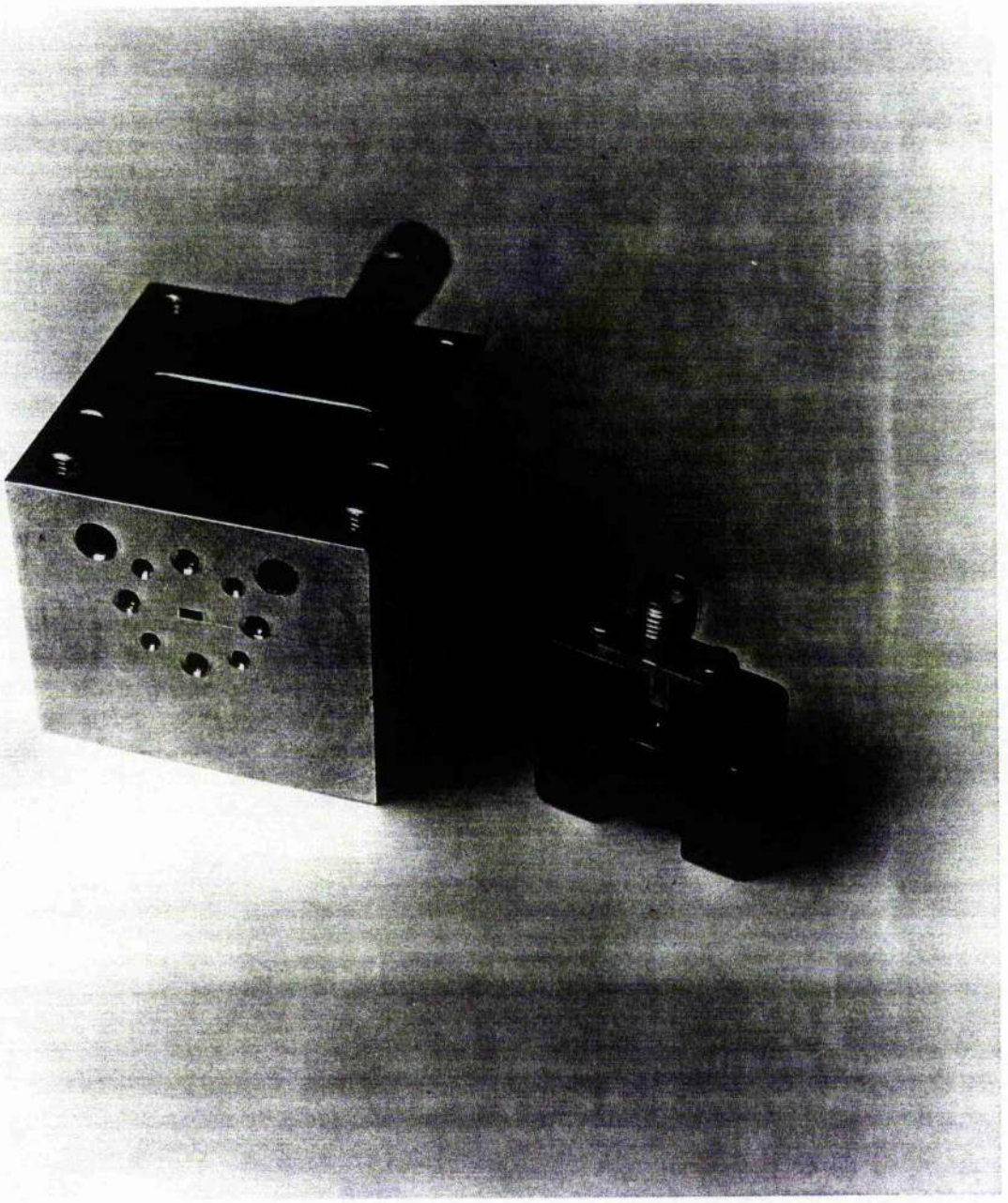


Fig. 2·14 : Detector Wafer



Detector Block and Wafer Insert

REFERENCES

[1] Smith, G.M., *Transferred Electron Oscillators at MM Wave Frequencies and their characterisation using Quasi-Optical Techniques*, Ph.D. Thesis, St. Andrews, March 1990.

[2] Carlström, J.E. et al, "A Continuously Tunable 65-115GHz Gunn Oscillator", *IEEE Trans Microwave Theory & Techniques*, **MTT-33** (7), July 1985, pp. 610-619.

[3] Webb, M.R., *Millimetre Wave Quasi-Optical Signal Processing Systems*, Ph.D. Thesis, St. Andrews, July 1992.

[4] Webb, M.R., "A millimetre-wave pyroelectric detector", *Intl. Jnl. IR & MMWaves*, **12** (10), Oct. 91, p. 1225.

[5] Leeson, M.L., *The Application of Quasi-Optical Techniques to Millimetre Wave Radar*, Ph.D. Thesis, St. Andrews, Aug. 1992.

[6] Precision Monolithics Inc., SSM-2100 Monolithic Log/Antilog Amplifier.

[7] Harvey, A.R., *A Millimetre-Wave Quasi-Optical Complex Impedance Bridge*, Ph.D. Thesis, St. Andrews, Aug. 1990.

[8] Thomas Keating Ltd., Station Mills, Billingshurst, West Sussex, RH14 9SH.

[9] National Semiconductor Corp., MF10 Universal Monolithic Dual Switched Capacitor Filter.

[10] Martin, D.H., "Polarizing (Martin-Puplett) Interferometric Spectrometers for the Near- and Submillimeter Spectra", *Infrared and Millimeter Waves*, Ed. K.J. Button., **6**, Ch. 2, Academic Press, 1982.

[11] Lesurf, J., *Millimetre-Wave Optics, Devices and Systems*, Ch. 9, Adam Hilger/IOP, 1990.

[12] Wild Vision, 15 Witney Way, Boldon Business Park, Boldon Colliery, Tyne & Wear NE35 9PE.

[13] Goodwin Electronic Measuring Systems, Poole Hall Ind. Est., Ellesmere Port, Cheshire L66 1ST.

[14] Farran Technology Ltd., Ballincollig, Cork, Ireland.

[15] Sharpless, W.M., "Wafer-Type Millimeter Wave Rectifiers", *B.S.T.J.*, Nov. 1956, pp. 1385-1402.

Chapter Three
Frequency Modulation and Power Supply design

Frequency modulation (FM) transmitters are based on a device called a modulator which varies the frequency of a carrier in sympathy with a modulating signal. Gunn oscillators can be used as voltage controlled oscillators (VCO) and as such are suitable as FM modulators. In this chapter I will present some basic FM theory and look at the requirements of power supplies for use with a frequency modulated Gunn diode. The FM behaviour of the Gunn device itself will be treated in the next chapter.

3.1 Fundamentals of FM

3.1.1 Definitions

A system which exhibits angle modulation can be described most generally as one in which information is transmitted by varying the phase angle of a carrier wave. Usually the carrier is assumed to be sinusoidal and can be represented as

$$s(t) = A\cos\{\theta(t)\} \qquad 3.1$$

where $\theta(t)$ is the carrier phase angle and varies in relation to the modulating signal $m(t)$. If $\theta(t)$ varies linearly with the amplitude of the modulating signal then this regime is known as phase modulation (PM). Defining the instantaneous frequency f_i as the rate of change of phase *i.e.*

$$\dot{\theta}(t) = 2\pi f_i \qquad 3.2$$

then frequency modulation (FM) is produced when $\dot{\theta}(t)$ varies linearly with the amplitude of the modulating signal.

Writing the carrier phase angle for PM as

$$\theta(t) = 2\pi f_c t + \varphi_o + K_1 m(t) \qquad 3.3$$

where f_c is the carrier frequency, φ_o is the initial phase and K_1 is a

constant, then given that

$$\dot{\theta}(t) = 2\pi f_i = 2\pi f_c + 2\pi K_2 m(t) \qquad 3.4$$

where K_2 is a constant, the corresponding expression for FM is

$$\theta(t) = 2\pi f_c t + \varphi_o + 2\pi K_2 \int_0^t m(t)\,dt \qquad 3.5$$

Comparing equations 3.3 and 3.5 one can see that FM is a special case of PM in which the carrier phase angle is linearly proportional to the integral of the modulating signal.

3.1.2 Frequency Deviation & Modulation Index

The frequency deviation of an FM signal is defined as the maximum departure of the instantaneous frequency from the carrier frequency. If the modulating signal is $m(t) = \beta \sin(\omega_m t)$, where β is the peak amplitude and $\omega_m = 2\pi f_m$, then the modulated carrier is

$$s(t) = A \cos\{\omega_c t + \beta \sin(\omega_m t)\} \qquad 3.6$$

β is the maximum phase deviation of the modulated carrier and is known as the modulation index.

Now from from 3.2

$$f_i = \frac{1}{2\pi}\dot{\theta}(t) = f_c + \beta f_m \cos(\omega_m t) \qquad 3.7$$

so the maximum frequency deviation, Δf, occurs when the cosine term is unity and is given by

$$\Delta f = \beta f_m \qquad 3.8$$

So the modulated carrier can be written as

$$s(t) = A \cos\left\{\omega_c t + \frac{\Delta f}{f_m} \sin(\omega_m t)\right\} \qquad 3.9$$

Note that while the instantaneous frequency lies in the range $f_c \pm \Delta f$, the signal spectrum contains components outside this range.

3.1.3 The spectrum of an FM signal — sinusoidal modulation

Consider the simple case of a carrier wave modulated by a

sinusoidal signal:

$$s(t) = A\cos\{\omega_c t + \beta\sin(\omega_m t)\} \qquad 3.10$$

This can be expanded in terms of a series of Bessel functions of the first kind to give[1]

$$\begin{aligned} s(t) = A\,[&J_o(\beta)\cos(\omega_c t) \\ &- J_1(\beta)\{\cos(\omega_c - \omega_m)t - \cos(\omega_c + \omega_m)t\} \\ &+ J_2(\beta)\{\cos(\omega_c - 2\omega_m)t + \cos(\omega_c + 2\omega_m)t\} \\ &- J_3(\beta)\{\cos(\omega_c - 3\omega_m)t - \cos(\omega_c + 3\omega_m)t\} \\ &+ \ldots] \end{aligned} \qquad 3.11$$

Using the following two properties of Bessel functions

$$J_n(\beta) = (-1)^n J_{-n}(\beta) \qquad 3.12$$

$$\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1 \qquad 3.13$$

one arrives at the general expression for a sinusoidally modulated cosine carrier:

$$x(t) = A\sum_{n=-\infty}^{\infty} J_n(\beta)\cos(\omega_c + n\omega_m)t \qquad 3.14$$

The spectrum of this signal can be determined by inspection - it has a component at the carrier frequency and an infinite number of sidebands on either side of the carrier, spaced at integer multiples of the modulation frequency. The number of sidebands increases with the value of β, but in a nonlinear manner, through the coefficients of the Bessel functions. This nonlinearity between the time and frequency domains marks one of the principal differences between frequency and amplitude modulation (AM). Expression 3.12 means the symmetric components about f_c are 180° out of phase and 3.13 ensures that the signal power is normalised over the whole spectrum such that power in the sidebands may appear only at the expense of power originally in the carrier.

3.1.4 The bandwidth of an FM signal

Whilst in principle there are an infinite number of sidebands in an FM spectrum, implying the bandwidth is also infinite, in practice, the majority of the signal power is confined to a finite number of sidebands in a finite bandwidth. The spectrum of an FM signal can be defined as containing all sidebands components with an amplitude $\geqslant$ 1% of the unmodulated carrier amplitude[2].

For sinusoidal FM, the nominal bandwidth is given by Carson's Rule:

$$B \cong 2(\Delta f + f_m) \qquad 3.15$$

Two limiting cases can be formed from Carson's Rule - for $\beta \ll 1$ the bandwidth is $B \approx 2f_m$ and this is called Narrow Band FM (NBFM) and for $\beta \gg 1$ the bandwidth is $B \approx 2\Delta f$, and this regime is known as Wide Band FM (WBFM). Carson's Rule actually works for any form of modulating signal which is band limited and has finite power[3].

3.1.5 The spectrum of an Wide Band FM signal

For a general pulse waveform (wide bandwidth) one can specify the peak frequency deviation $\Delta\omega$ and the time duration T. If the waveform is periodic then T is the period and the dispersion index is defined as[4]

$$\beta_1 \equiv \left(\frac{\Delta\omega}{2\pi}\right)T \qquad 3.16$$

and this is used in place of the modulation index.

For very low dispersion indices the spectral content of the modulating signal largely determines the magnitude of the FM spectral density. For very high dispersion indices, the amplitude-to-frequency conversion largely determines the magnitude spectral density.

3.1.6 Woodward's Theorem

If one lets $\beta \rightarrow \infty$ one would expect the amplitude-to-frequency conversion to be dominant. In terms of spectral density, this means that the spectral magnitudes are in proportion to the fractional time spent at each frequency.

Taking the FM signal $s(t) = A\cos(\omega_m t)$ as an example, then the frequency deviation about the carrier, $\omega_d = \omega_i - \omega_c$, is given by

$$\omega_d = \Delta\omega \cos(\omega_m t) \qquad 3.17$$

Rearranging for t gives

$$t = \frac{1}{\omega_m} \cos^{-1}\left(\frac{\omega_d}{\Delta\omega}\right) \quad ; \text{for } |\omega_d| \leqslant \Delta\omega \qquad 3.18$$

The fractional amount of time per unit frequency is then

$$\frac{1}{T}\left|\frac{dt}{d\omega_d}\right| = \frac{1/2\pi}{\sqrt{1-\left(\frac{\omega_d}{\Delta\omega}\right)^2}} \quad ; \text{for } |\omega_d| \leqslant \Delta\omega \qquad 3.19$$

and this is the probability density function of the modulating waveform. This function is shown in Fig. 3·1.

This property is described by Woodward's Theorem which states that "The power spectral density of a WBFM waveform is approximated by, and has the same form as, the probability density function of the modulating waveform".

Blachman and McAlpine[5] analyse the validity of this approximation for the case of FM by Gaussian noise. They conclude that for any real system which includes an input filter (that effectively smears out the finer details of the spectrum), Woodward's Theorem is a useful approximation for FM spectra when the modulation index is large.

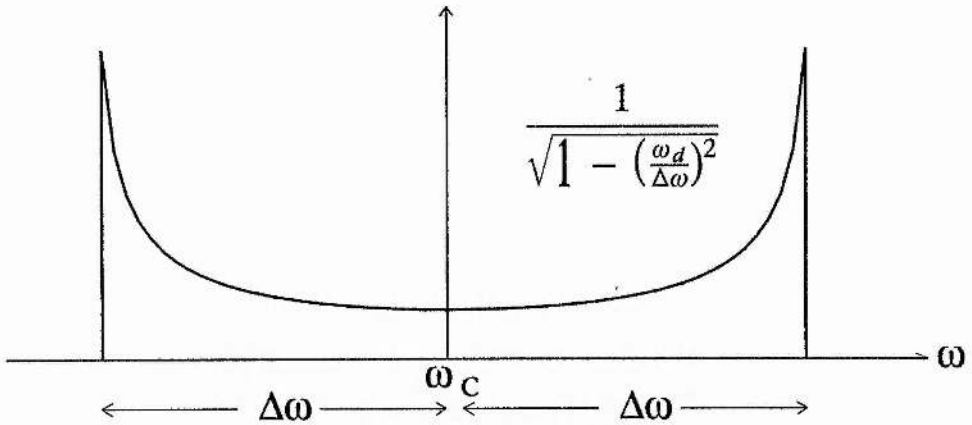


Fig. 3·1 : "Crown" spectrum of wide band sinusoidal FM

3.2 Power Supply Design

3.2.1 Power Supply Requirements

The quality of a Gunn oscillator power supply unit (PSU) is a determining factor in the performance of the oscillator. Primarily, the bias voltage must be stable and have a very low level of noise since bias fluctuations will contribute to frequency instability and phase noise. Additionally, the output impedance of the PSU must be low to minimise oscillations on the bias line between the supply and the diode.

The bias voltage needs to be smoothly variable from zero up to about 6V for GaAs diodes and up to about 11V for InP devices. These figures are maximum values and in practice operation would be at lower bias voltages. Current supply needs to be up to 1.5A and 250mA for GaAs and InP diodes respectively.

In order that a Gunn oscillator can be frequency modulated, the power supply must have provision for a modulation input. Any external modulation signal is then superimposed on the bias voltage.

The following two sections describe two types of power supply which have been used in this work.

3.2.2 Type 1 PSU

The standard type of power supply I use is based round discrete transistors running off a single supply rail. The unit itself takes power from a 30V 3A lab bench supply. A regulator IC is used to derive the bias voltage and a series pass power transistor (emitter follower) is used to provide the current drive. The circuit appears in Fig. 3·2 overleaf.

The circuit operates in the following manner. Bias is derived from the 7812 regulator and is adjustable by means of the 10-turn pot VR_1. Since a maximum of 11V is required and there are two diode drops due to the transistors, R_1 and D_1-D_5 are included to raise the output of the regulator to above its nominal 12 volt rating. The output pin sits 12V above the centre pin which is held at ~2.4V and the input pin must be at least 2.5V above the output so the input voltage must be at least 17V. C_1-C_6 are all decoupling capacitors.

The modulation input is based round Q_1, a common-emitter amplifier. R_6 is used to match 50Ω cables and R_4 and R_5 are a potential divider, set empirically to counteract the gain of the transistor to ensure an overall modulation gain of unity. C_8 limits the modulation bandwidth and D_1 prevents the output from swinging negative at low bias voltages. VR_2 is used to partially turn on Q_1 to adjust its operating point and ensure maximum linear modulation. Note that the modulation input is inverting.

The bias and modulating signals are added together at Q_2 via R_2 and R_3 with C_7 removing noise. Q_2 and Q_3 provide the high current drive at a given bias voltage necessary for supplying the Gunn diode. R_8 is a current limiting resistor, as is R_9 on the monitor socket. R_8 also lowers the voltage dropped across Q_3 to reduce its static dissipation. The voltmeter is set up to give the required full-scale deflection.

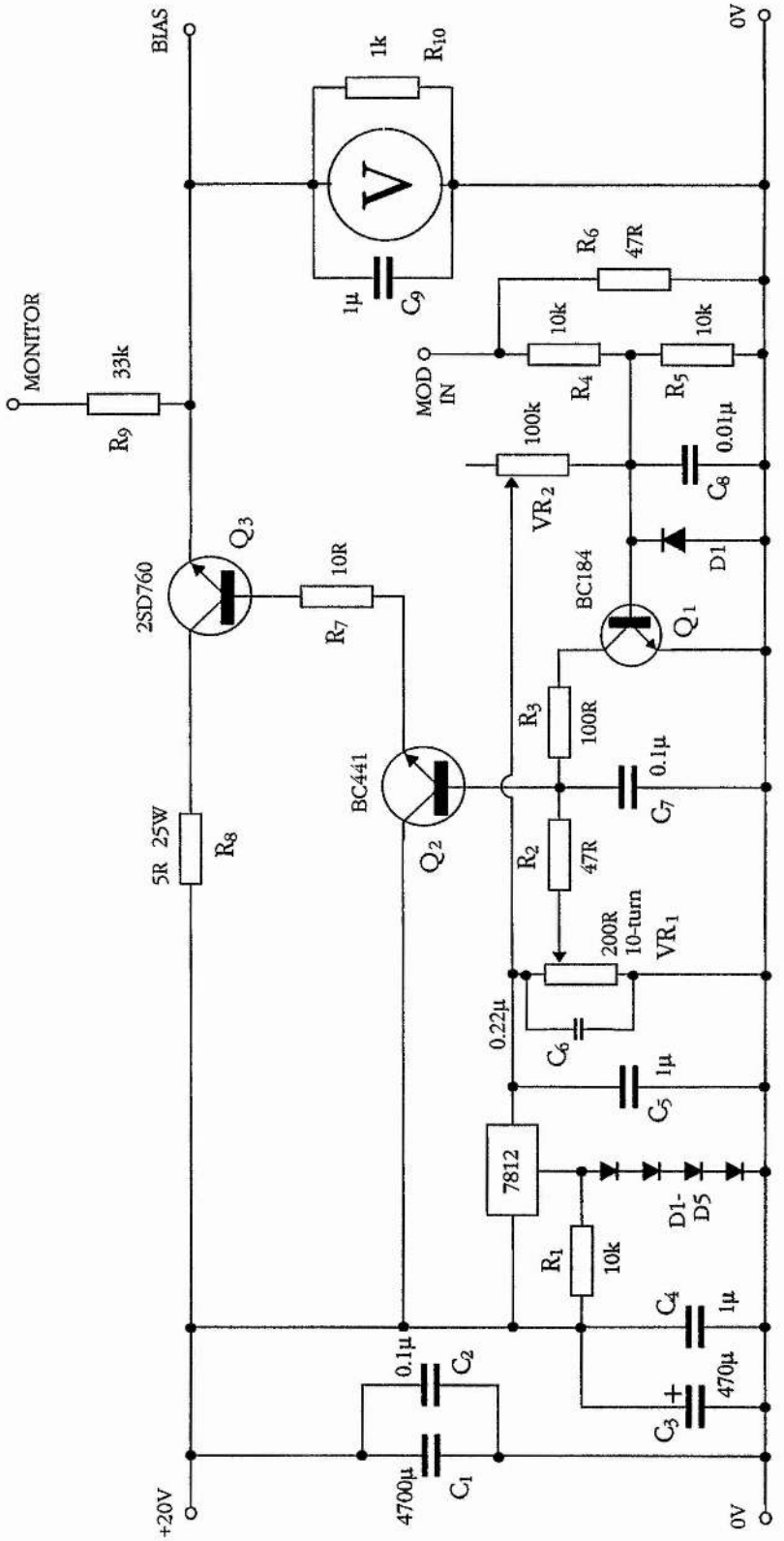


Fig. 3·2 : Type 1 Power Supply Circuit

3.2.3 Type 2 PSU

To be able to modulate Gunn diodes at much higher frequencies than achievable with the Type 1 PSUs, I designed a new circuit which is inherently very wideband. The new PSU is based on integrated circuits rather than discrete devices and each IC was chosen for high speed and low noise performance. The design is based around an Elantec 2009CT current buffer amp which has a 90MHz bandwidth and a current rating of one amp.

The circuit is shown overleaf. The bias is derived from a regulator as before and is filtered by R_1 and C_7 before being buffered by IC_1, a low noise OP-90G. Modulating signals are buffered by an OPA-633KP chosen for its full power bandwidth of 40MHz. The resistors R_4 & R_5 are included to limit the supply current drawn by IC_2 and hence reduces its quiescent power dissipation.

The bias and modulation are combined in a summing amplifier, IC_3, but since the chosen device (an LM6364N) is only stable for gains greater than five, a potential divider (R_6 & R_7 with R_8) is included to compensate.

The series diode preceding IC_3 prevents the output swinging negative if a modulating signal is present but the bias is at zero. VR_2 is used to trim out any dc offset at the output of the supply. The current drive is provided by IC_4 and the monitor socket includes a 100Ω current limiting resistor. Many decoupling capacitors are used in the circuit as recommended in the IC data sheets and the whole circuit is built on a double sided PCB, with the top side connected as a ground plane. These measures ensure circuit stability and minimise leakage at high frequencies.

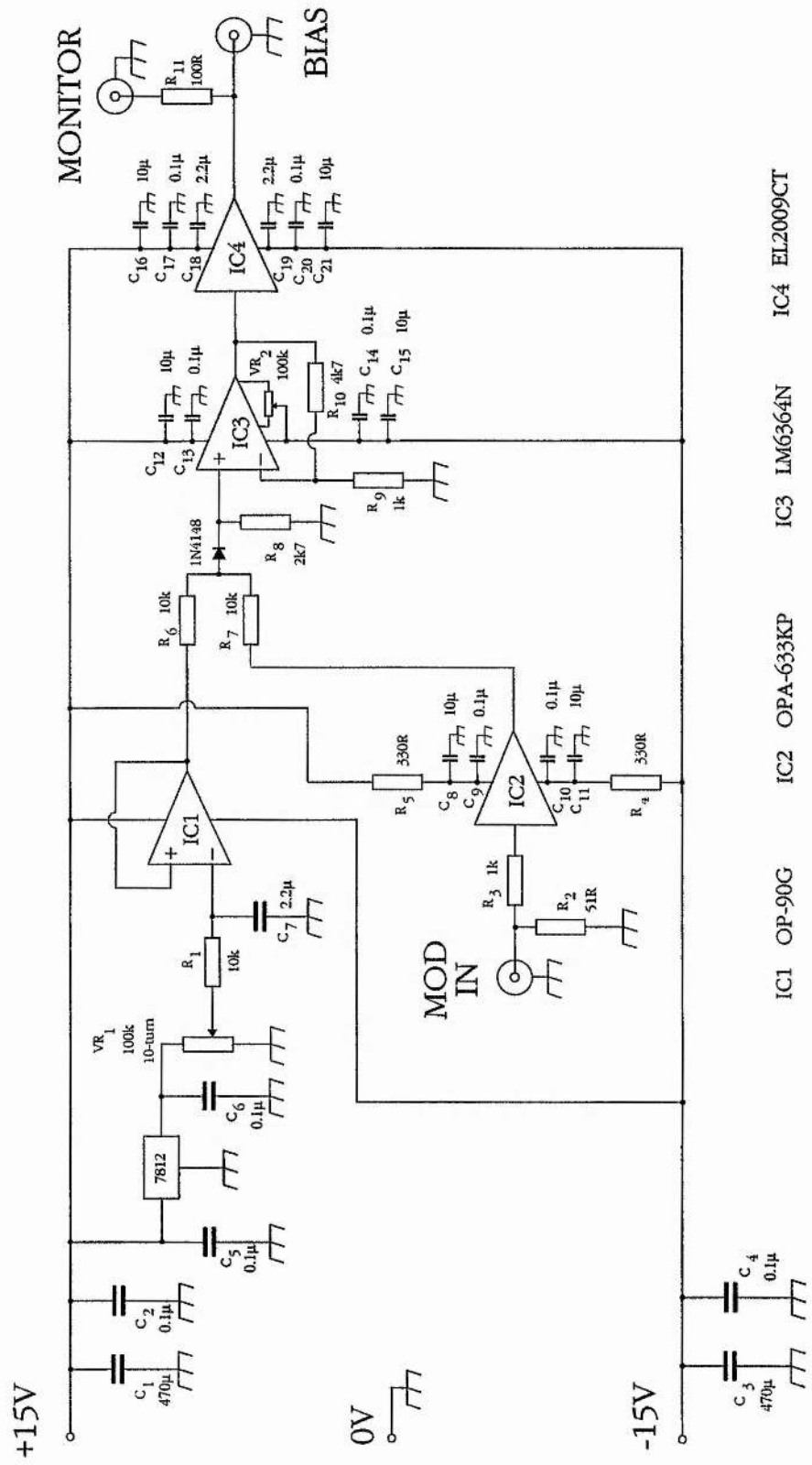


Fig. 3·3 : Type 2 Power Supply Circuit

3.3 Power Supply Results

3.3.1 Type 1 PSU

The performance of the Type 1 power supply is a trade-off between low noise and the modulation bandwidth controlled by C_7 & C_8. The circuit has an inherent maximum bandwidth of about 450kHz but in practice this is limited to 100kHz or less. The lowest noise version of this PSU has a modulation bandwidth of only 200Hz. Note that the series resistor on the monitor socket, coupled with cable capacitance can roll-off high frequencies, so for a wideband supply the value of this resistor has to be reduced. The output impedance of this type of supply was measured as approximately 0.6Ω for voltages above threshold (*i.e.* >1V).

3.3.2 Type 2 PSU

The integrated circuit design of power supply (Type 2) has a flat frequency response up to 10MHz but 3dB of gain peaking is evident around 20MHz. See Fig. 3·4 overleaf.

The bandwidth has not been deliberately limited in this design and Fig. 3·4 reflects the inherent bandwidth of the circuit. A small amount of input capacitance may reduce the gain peaking and extend the usable (flat) response to higher frequencies. The gain does not vary with bias voltage and the supply can drive either GaAs or InP diodes. The output impedance of the EL2009CT is quoted as 1 ohm[6].

Certain operating precautions need to be observed when using this PSU. Any output offset voltage should be trimmed to a negligible level, and the unit must be switched on or off with the Gunn diode disconnected. Since this is a dual supply rail design with an IC in the output, when the circuit is switched on or off, the output has a tendency to jump to one of the supply rails, which would be catastrophic for the Gunn diode.

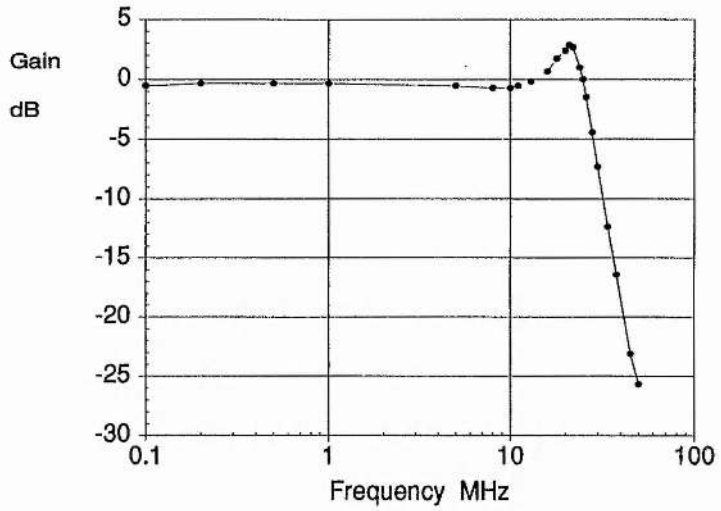


Fig. 3·4 : Type 2 PSU Frequency Response

3.3.3 Comparative Noise Performance

To assess the relative noise performance of different power supplies the noise power spectral density (NPSD) of each was measured from DC to 100kHz with a Stanford Research Systems SRS760 FFT Analyser. The SRS760 gives a direct reading of NPSD in $V/\sqrt{Hz}$ and measurements were made for four PSUs driving about 0.9A into a resistive load at 5V. Three of the units were of Type 1 design and had bandwidths of 200Hz, 20kHz and 100kHz. The fourth unit was the Type 2 design.

Most of the voltage noise on the bias supply of an oscillator will appear as FM noise on the output signal due to the voltage-to-frequency conversion effects in Gunn diodes. Voltage-to-AM noise is a smaller effect and shall not be considered here.

The amplitude of the first FM sideband is proportional to the Bessel component $J_1(\beta)$ which is approximately $\beta/2$ for small β. [N.B. $\beta = \Delta f / f_m$ and $\Delta f \propto A_m$]. The condition that β be small is met in this case since A_m, the noise voltage, is very small and hence Δf is also small. The condition breaks down only at very low values of f_m (i.e. near DC).

Therefore, substituting β by $\Delta f / f_m$ and Δf by A_m, the relative FM noise amplitude is proportional to A_m / f_m. So by dividing the PSD by the frequency one obtains a comparative measure of the noise performance of different power supplies. The results of these measurements are shown in Fig. 3·5.

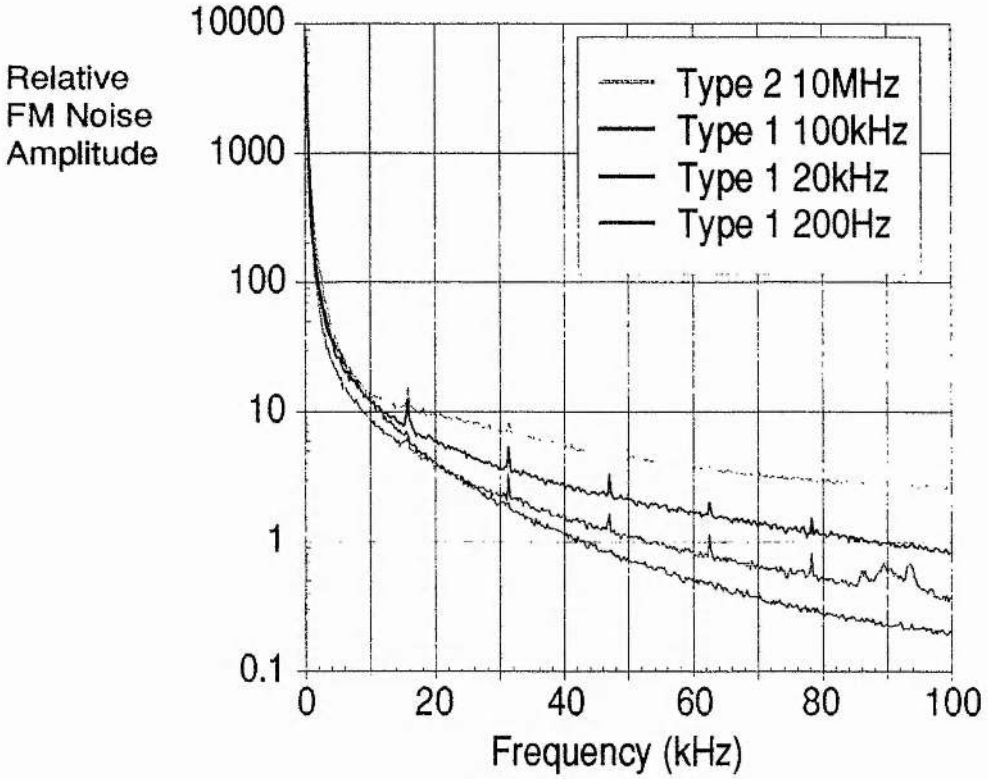


Fig. 3·5 : Comparative Noise Performance of Power Supplies

N.B. The spurious peaks are caused by extraneous interference and are not a feature of any of the PSUs.

It can be clearly seen that there is a correlation between bandwidth and noise contribution. In practice, I have found that the 200Hz Type 1 supply has the lowest noise and the Type 2 unit is noticeably worse. Since the FM sideband amplitude is inversely proportional to the modulating frequency, it is low frequency noise that has the greatest effect on an oscillator. Whilst the Type 2 design has a considerably greater high frequency noise floor (over many tens of megahertz) compared with the Type 1 circuit (in which the

transistors have an f_T of no more than 1MHz) the low frequency noise contributions are not dissimilar. Hence, careful attention must be paid to minimising noise on the regulator/bias stage of both types of supply.

I attempted to reduce the level of high frequency noise in the Type 2 supply with various modifications to the circuit. The majority of the noise was traced to the LM6364N op-amp and I found that much lower noise was achievable by substituting an OP-27. However, this device has insufficient bandwidth and I replaced it with an LM6361N. This IC has a higher quoted noise than the 6364 but by lowering the values of R_9 and R_{10} I was able to make a significant improvement over the original circuit. Fig. 3·6 shows the noise spectrum of each configuration. The measurements were taken with a Hewlett-Packard HP3585B spectrum analyser.

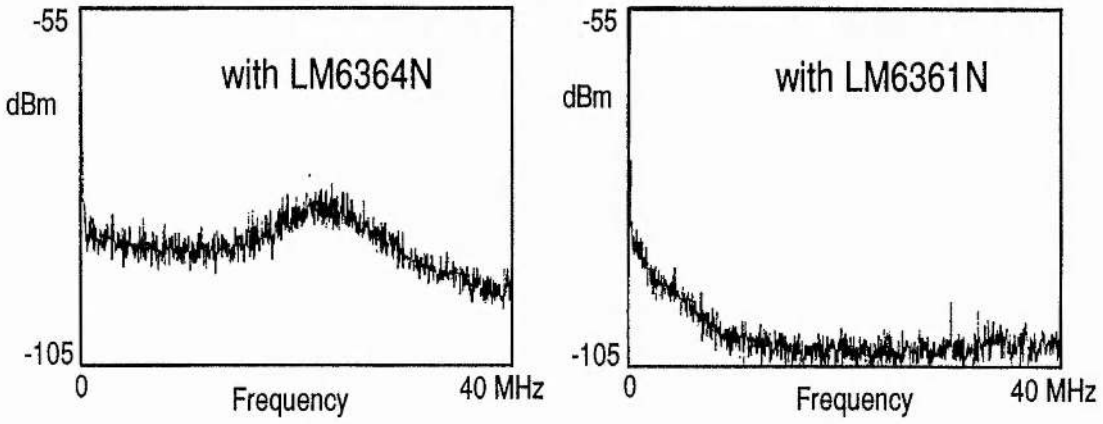


Fig. 3·6 : Noise spectra for Type 2 PSU with LM6364N & LM6361N

Regrettably, the lower noise variant suffered from reduced bandwidth and distortion of the modulation for amplitudes greater than $0.8V_{rms}$. In view of this, the circuit was returned to the original form as shown in Fig. 3·3 and I have accepted the compromise between noise and bandwidth.

An obvious qualitative comparison of power supply noise can be made when observing the spectrum of an oscillator, downconverted with a heterodyne receiver. The width of the IF

peak relates to the noise performance of the power supplies being used for the signal and LO oscillators. I made comparisons of several power supplies in this way and Fig. 3·7 shows the best and worst units. The lowest noise was achieved with the 200Hz bandwidth Type 1 supply, as mentioned above, and the highest noise supply was the Type 2. However, in view of the difference in bandwidth of five orders of magnitude, I think the Type 2 design is quite respectable in terms of noise.

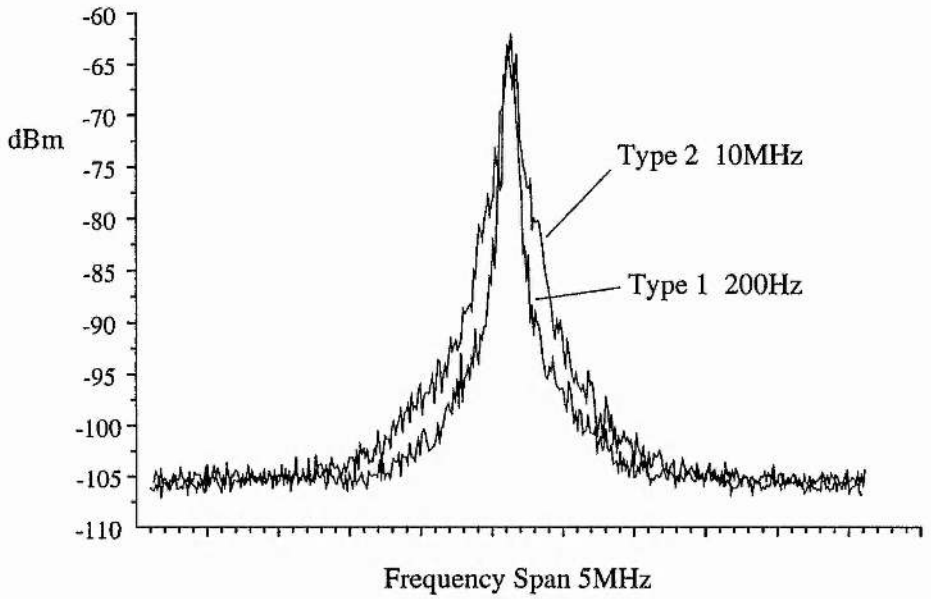


Fig. 3·7 : Increase in IF linewidth due to PSU noise

3.4 FM Spectra of Gunn Oscillators

Gunn oscillators work quite well as VCOs but their behaviour as FM modulators is not ideal. This is primarily due to the nonlinear variations of frequency and power as the device bias is varied. For NBFM applications, this is not a problem but as the modulation index is increased the nonlinearities become evident.

The bias tuning behaviour of Gunn diodes will be dealt with in detail in the next chapter but I will end here by showing some examples of FM Gunn oscillator spectra.

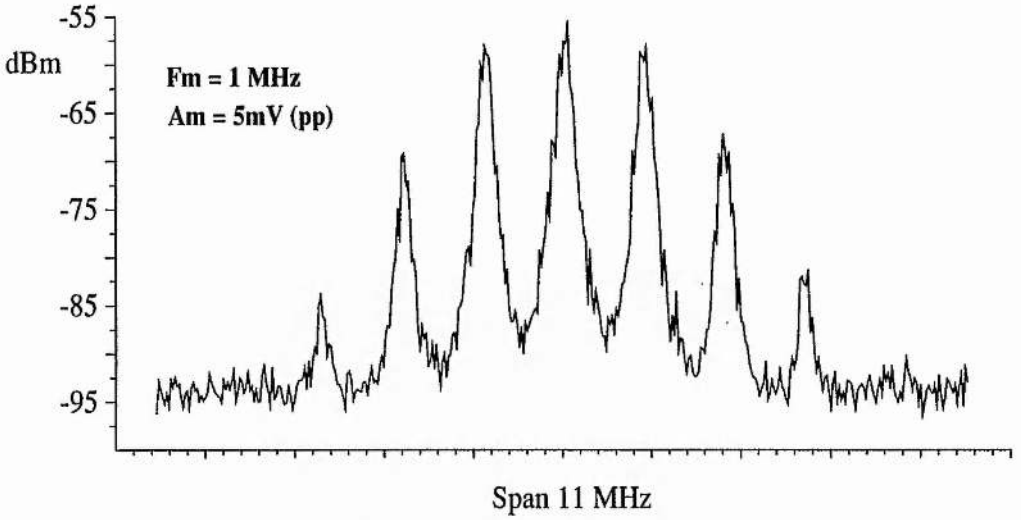


Fig. 3·8 : Example of NBFM spectrum

Above one can see a good example of classic low-β frequency modulation. No video averaging was used in any of these three traces, hence the rather jagged appearance.

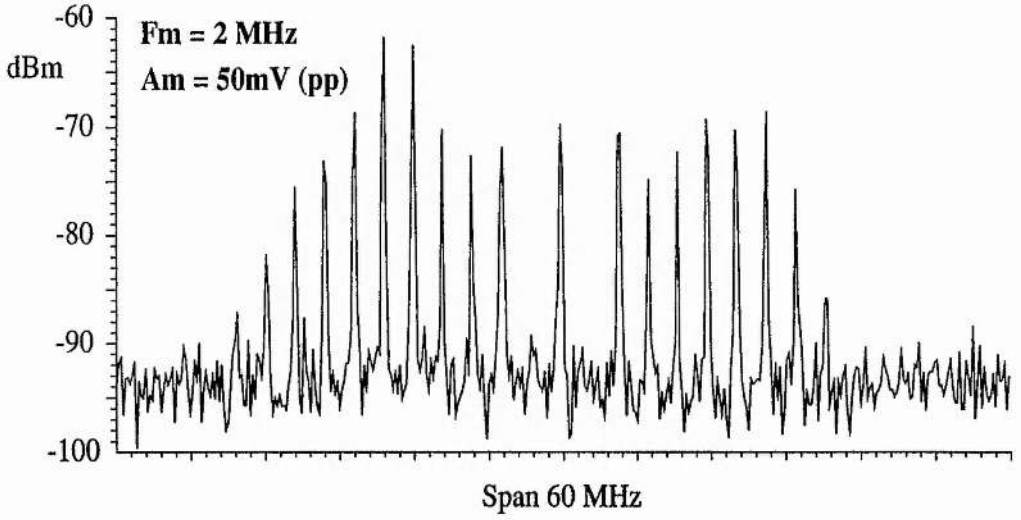


Fig. 3·9 : Distorted WBFM spectrum

Fig. 3·9 indicates the nonlinear modulation effects of the Gunn diode. Note the asymmetric amplitude of the envelope and the missing harmonics.

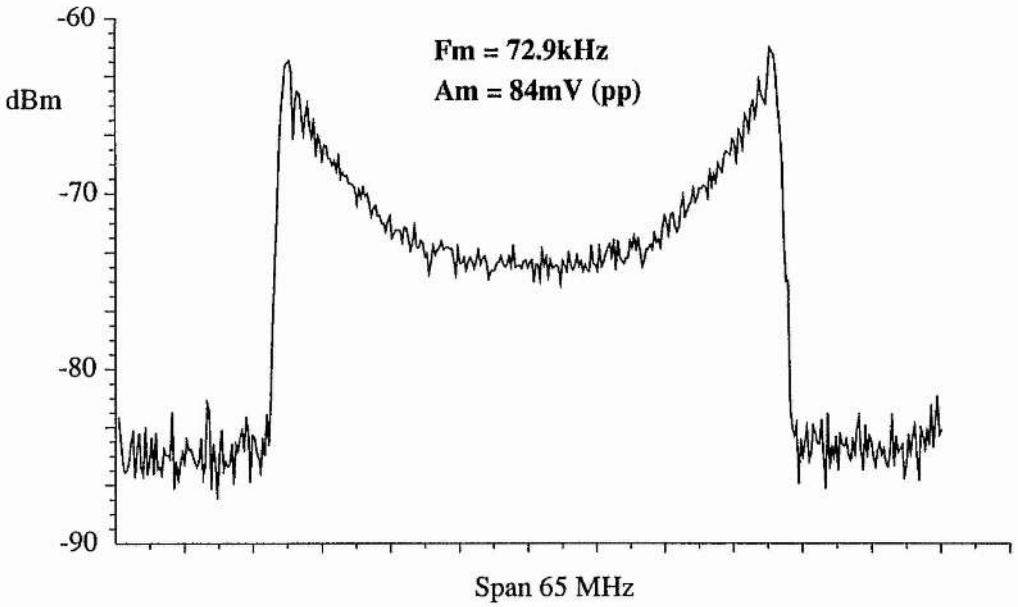


Fig. 3·10 : Classic "crown" spectrum of sinusoidal WBFM

This shows very good comparison with the theoretical result shown in Fig. 3·1.

REFERENCES

[1] Taub, H. & Schilling, D.L., *Principles of Communication Systems*, McGraw-Hill, 1971.
[2] Dunlop, J., & Smith, D.G., *Telecommunications Engineering*, Chapman & Hall, 2nd Ed., 1989, p. 69.
[3] Stremler, F.G., *Introduction to Communication Systems*, Addison-Wesley, 1977, p. 277.
[4] ibid., p. 270.
[5] Blachman, N.M., & McAlpine, G.A., "The Spectrum of a High-Index FM Waveform: Woodward's Theorem Revisited", *IEEE Trans. Commun. Tech.*, **COM-17** (2), April 1969, pp. 201-208.
[6] Data Sheet, Elantec Inc., 1996 Tarob Court, Milpitas, CA, USA.

Chapter Four
Bias Tuning Measurements

In applications where Gunn oscillators are used as frequency modulated sources, it is necessary to know the modulation characteristics of the Gunn diode utilised. Assuming the modulation is imposed via the bias voltage to the Gunn, then one needs to measure the frequency and power versus applied bias. Modulation can also be achieved by incorporating a varactor diode into the Gunn oscillator circuit, thereby providing a voltage dependent tuning capacitance, but I will not discuss this scheme here.

Assuming that the power supply driving the oscillator is linear, an ideal frequency modulator would have a linear bias/frequency characteristic and a flat power response over the range of bias voltages used. In practice, Gunn diodes exhibit nonlinearity in both parameters and it is this behaviour which must be measured. This chapter examines different ways of measuring the bias tuning characteristics of Gunn diodes and discusses some of the mechanisms responsible for these voltage dependent properties.

4.1 Manual Method

The traditional, manual method for measuring the bias tuning characteristics of a Gunn oscillator involves stepping the bias voltage through a range of values and measuring the frequency and power for each setting. The frequency can either be measured with a microwave counter or by using an interferometer to calculate the signal wavelength, while the power is measured with a waveguide power meter. The experimental set up is the same as shown in Fig. 2·1.

Unfortunately, this method is slow and very tedious. Additionally, the time taken between bias increments tends not to be constant, due to human inexactitude, so significant thermal drift can

occur which tends to give inconsistent results. Hence, I devised an automated system to overcome these limitations.

4.2 Computer Controlled System

The computer controlled system which has been developed for making automatic bias tuning measurements of Gunn oscillators works on the following principles.

A digital-to-analog convertor (DAC) is stepped over a range of voltages under control of software. The output of the DAC is applied to the modulation input of a Gunn oscillator power supply in order to increment the bias over a given range. At each increment of voltage the computer pauses for a given time before recording the values of bias, frequency and power. Once the full sweep of bias voltages has been completed, the computer saves the data in a format suitable for exporting into a graph drawing package. Each part of the system will now be described in more detail. The system is shown in Fig. 4·1.

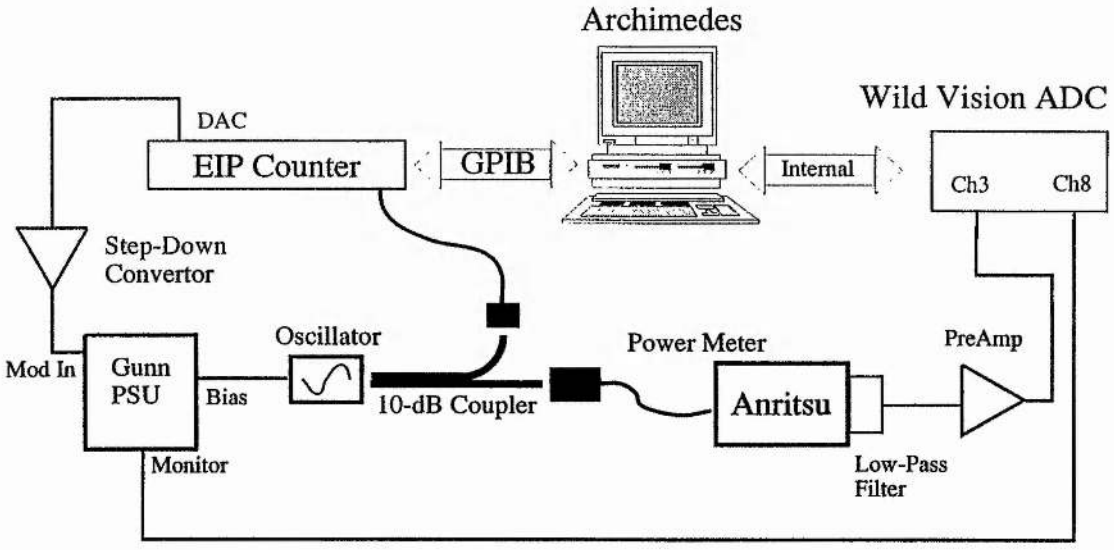


Fig. 4·1 : Automatic Bias Tuning System

4.2.1 The Controlling Computer

The computer used here is an Archimedes™ 410/1 with 4MB of memory and a hard disc. It is fitted with an Intelligent Interfaces IEEE.488 (GPIB) card[1] and a Wild Vision ADC1208-16 analog to digital converter (ADC) card[2]. The GPIB card is supplied with routines in BASIC which have been implemented in the controlling software. The Wild Vision card is an eight input 12-bit ADC with a maximum sampling rate of 167kHz. The input voltage range is nominally ±5V. Controlling the card from BASIC is simple and an on-board ROM holds all the necessary software modules.

4.2.2 The DAC

The EIP 578 Source Locking Microwave Counter used in this system is equipped with a 12-bit DAC intended for frequency locking applications. It is possible to address the DAC directly and set its output to voltages in the range 0→+10V. At switch on, the DAC sits at +5V. By sending the necessary codes to the EIP over the GPIB interface, the computer is able to control the DAC and vary the bias applied to the oscillator under test.

Since the full 10V range of the DAC is too great to apply directly to the Gunn PSU, I added a step-down converter to the DAC output. This merely consists of a buffer amp (using a 77000 power op-amp) with a potential divider on the input. Therefore the full 12-bit range of the DAC can still be used but the voltage applied to the Gunn only sweeps over a volt or two, depending on the chosen step-down ratio. Typically, the bias is swept over about 1.5V.

4.2.3 Frequency Measurement

The EIP Counter is also used to measure the frequency of the oscillator. The counter uses a harmonic mixer[3], connected to the oscillator via a 10dB coupler, to measure the signal frequency with respect to a stable crystal reference. Frequency measurements are

updated every fraction of a second and the data is available via the GPIB interface. This is loaded directly into the computer.

4.2.4 Bias Voltage Measurement

The Wild Vision ADC card is used to measure the bias applied to the Gunn by sampling the output of the bias monitor socket on the PSU. Since the input range of the ADC is normally ± 5V and the bias voltage is often greater than this, I added a potential divider to one of the ADC inputs. The resistor values had to be chosen with care - too high a value would generate noise and limit the current into the ADC, and too low a value would load the signal source. In practice, I used a pair of 39kΩ resistors to extend the input range to ±10V. The value returned by the ADC is in the range 0 to 4095 so the nominal voltage reading is obtained from

$$\text{Voltage} = ((\text{ADC value} - 2048)/2048) \times 5.00$$

For the ±10V input, I calibrated the ADC against a digital voltmeter and found that the scale factor was not exactly 10.00 but actually 10.03. This is the value I have used in the code.

4.2.5 Power Measurement

An Anritsu ML83A power meter with a WG27 BiSb thermistor head is used to measure the oscillator output power. The instrument has a recorder output voltage which corresponds to the position of the analog meter on the front panel. This voltage is sampled with the ADC to record the power measurement.

The output voltage varies from 0→1V, with a source impedance of 1kΩ, but unfortunately it is not a direct representation of the measured power. It corresponds to the meter needle position which is autoranging, though this feature can be disabled to avoid jumps in range. Sampling the recorder voltage showed up large fluctuations on the dc level. This turned out to be 30mV of 440Hz

generated by the electronic chopping process implemented in the meter. In response to this, I added a low-pass filter with a time constant of 1.6s (1kΩ & 100μF) to the output socket, which reduced the ripple to a negligible level. The thermistor head has a response time of about two seconds so the filter does not restrict the rate at which measurements can be taken.

In order to make fuller use of the 5V input range of the ADC, a variable gain preamp is used to increase the power meter output. The gain is adjusted before each set of measurements and the software, described below, includes a routine to calibrate the sampled voltage in terms of milliwatts of power.

4.2.6 Software

The structure of the controlling software "Auto_Bias2" is outlined schematically below.

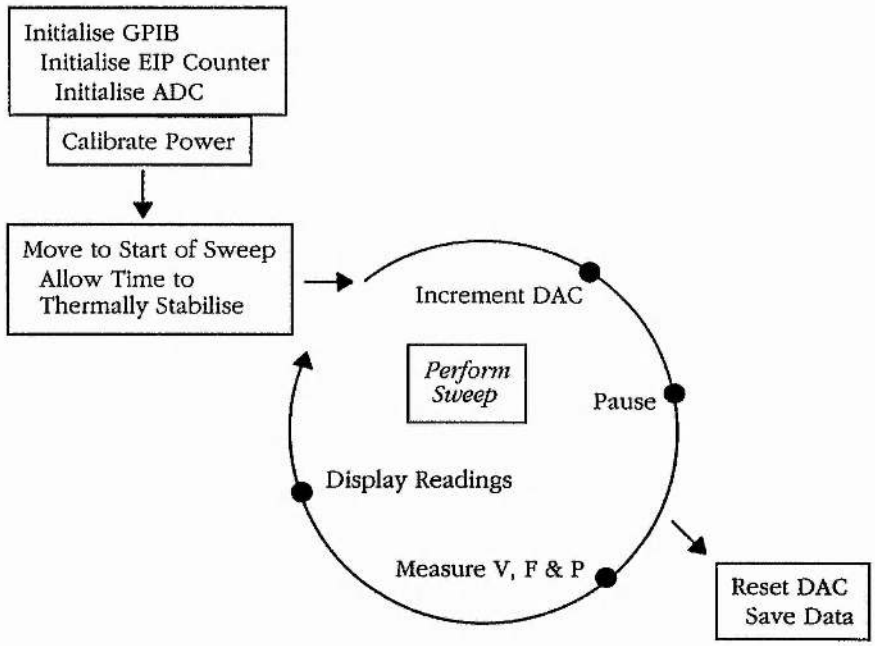


Fig. 4·2 : Auto_Bias2 Software Schematic

A full listing of the code, which is written in BASIC *V*, complete with detailed comments, is presented in Appendix 1.

4.2.7 Example Results

The computer controlled system has been a great success. It typically takes only 15-20 minutes to make a set of measurements on an oscillator which would previously have taken an hour or more. The bias voltage sweep can be adjusted by changing the step-down ratio of the DAC voltage and the number of points in a scan is easily changed in the software.

Additionally, because all the measured points are taken at equal time intervals, the consistency of scans is very good. Thermal drift is always the principal problem in this respect and computer control has greatly reduced this effect. Fig. 4·3 below shows three overlaid data sets which demonstrate the consistency of this technique. The frequency measurements are exceptionally good and the power readings show only a slight drift. The scans were taken about 15 minutes apart with an oscillator which had been left to warm up fully.

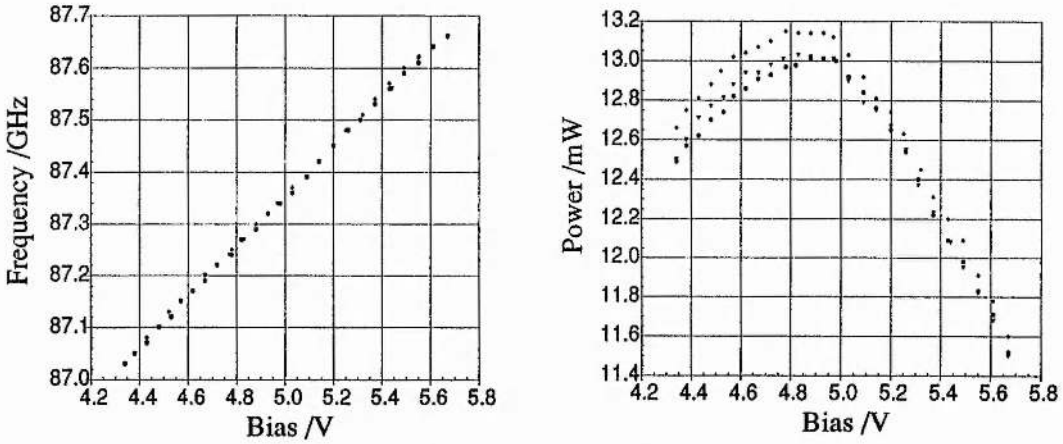


Fig. 4·3 : Example data from the Auto_Bias2 system

Being able to measure bias tuning characteristics quickly (and painlessly) has meant that I can make more detailed evaluations of this property of Gunn diodes. Results of such measurements are presented in the next section.

Finally, the system could be improved by using a different

power meter. A Boonton 4220 unit is now available which has a GPIB interface. This would eliminate the rather unsophisticated existing arrangement of analog-output/filter/preamp. Direct readings of frequency would be available via the GPIB and the software could be simplified by removing the power calibration routine. Unfortunately, lack of time has prevented me from making this upgrade.

4.3 Gunn Oscillator Measurements

The new automated technique for characterising oscillator bias tuning behaviour has meant that many data sets can be obtained easily, allowing detailed examination of tuning properties.

Using this system I was able to measure bias tuning curves for a particular oscillator over a whole range of oscillator settings. The results presented here are measurements taken on an experimental GEC diode DB637/6/1#1, mounted in a full-height WG-27 block (Block 1).

4.3.1 Choke and Backshort Varied

I measured the frequency and power tuning curves versus bias voltage, for a range of choke micrometer settings (cavity lengths). At each choke position, the backshort was adjusted to maximise the power.

As one might expect, the centre frequency of each scan changes monotonically with cavity length, with the higher frequencies corresponding to a shorter cavity. All the tuning curves are presented overleaf in Fig. 4·4. The choke micrometer positions were 1.30mm, 1.35, 1.40, 1.45, ..., 1.75, 1.80, 1.85*, 1.90, 2.00 & 2.05mm. Note the typical frequency jump in the curve fourth from the end (starred above).

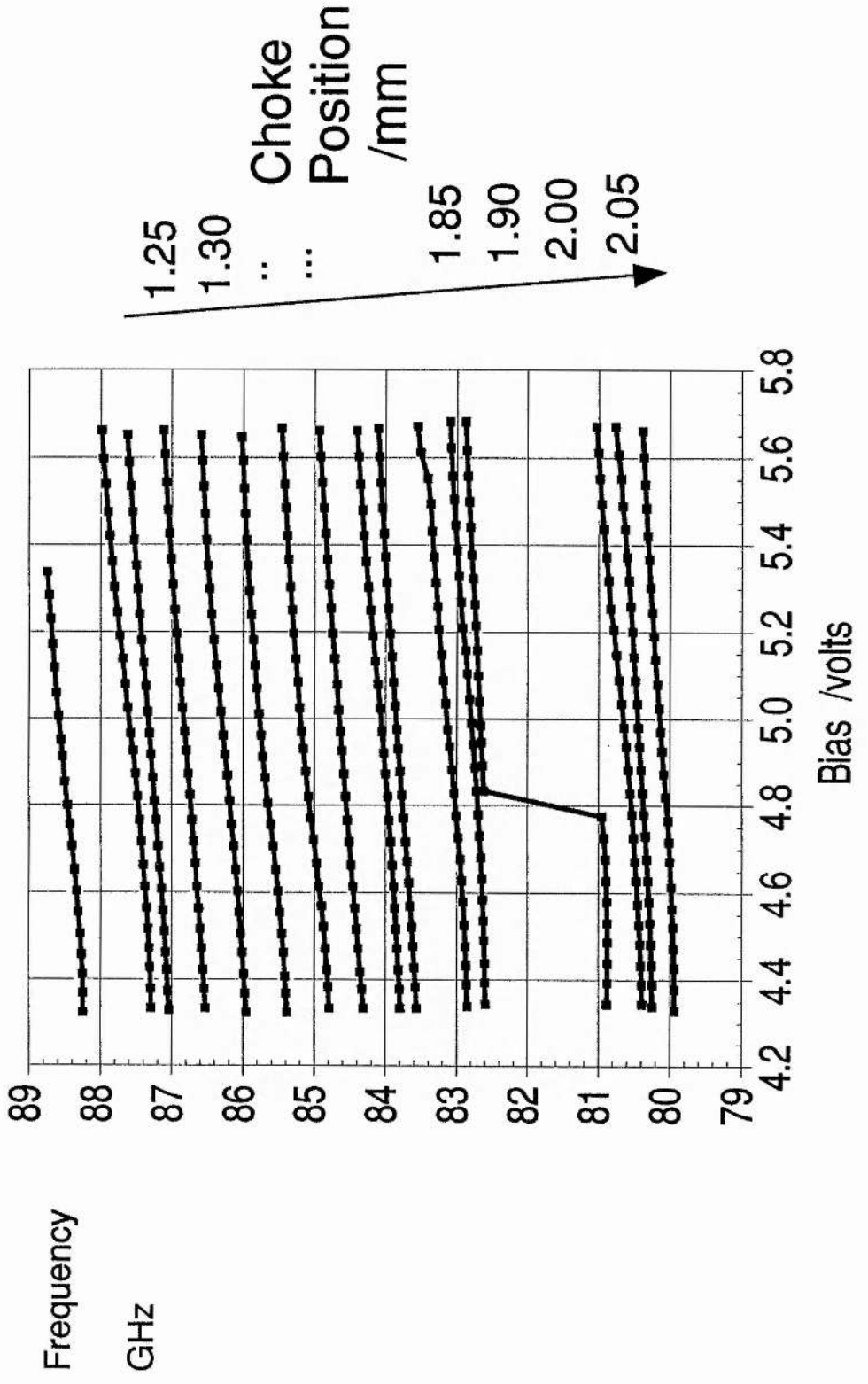


Fig. 4·4 : Tuning curves for range of cavity heights

The data as presented above do little to highlight the bias tuning behaviour. The total frequency tuning range of the oscillator is much greater than the voltage-frequency sweep so the gradient of each curve is suppressed. In order to avoid this effect, I have modified the data to show frequency offset with respect to a nominal reference point, taken as the frequency value at 5.0 volts bias. The bias is also displayed as an offset with respect to 5.0 volts. Fig. 4·5 shows the modified data with the omission of the curve containing the frequency jump.

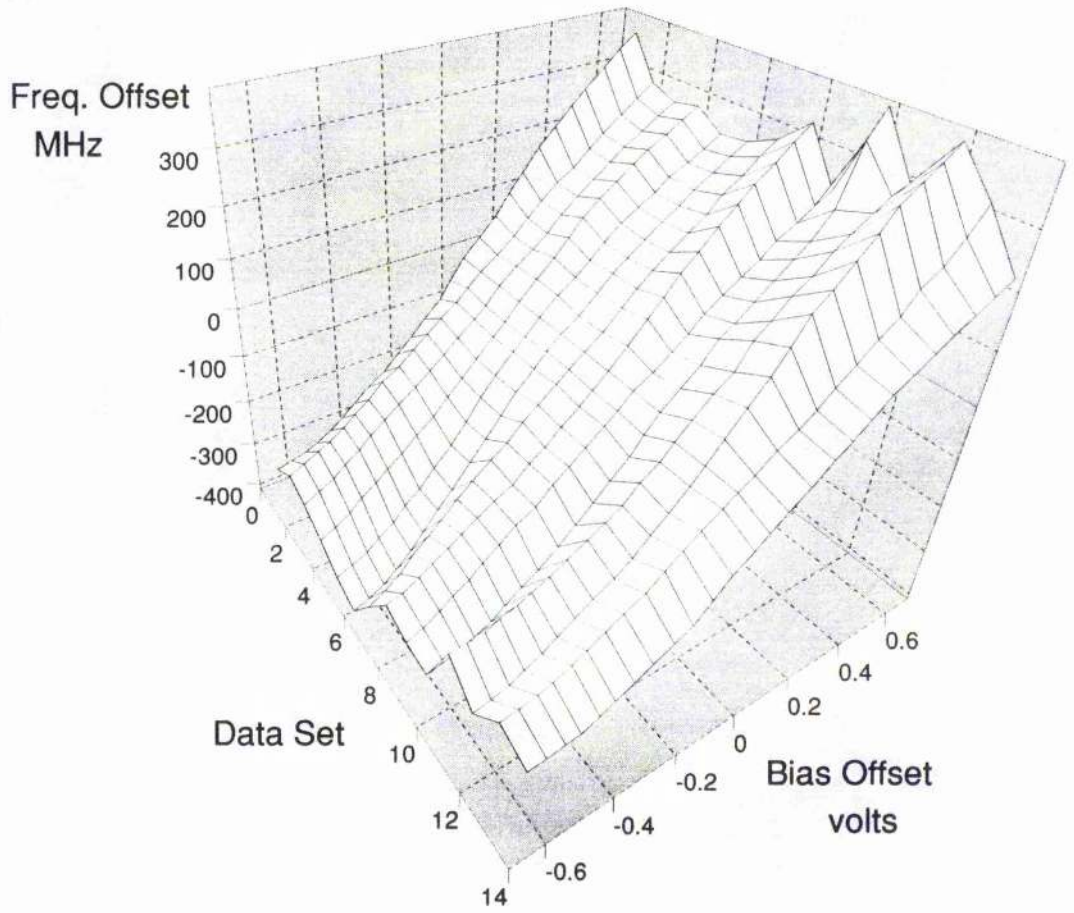


Fig. 4·5 : Frequency Tuning Curves

The set of tuning curves in Fig. 4·5 do not appear to show any particular trend. Since the data sets were obtained for oscillator settings which were different in both choke and backshort position, further measurements were made to isolate the two effects.

4.3.2 Backshort Varied - Choke Fixed

Tuning curves were recorded for the same oscillator whilst varying only the backshort position. The cavity length remained constant throughout. The frequency tuning results are presented below, again using the "offset from 5.0 volt reference" method. The backshort micrometer settings were 0.90mm, 1.00, 1.10, ..., 1.70mm.

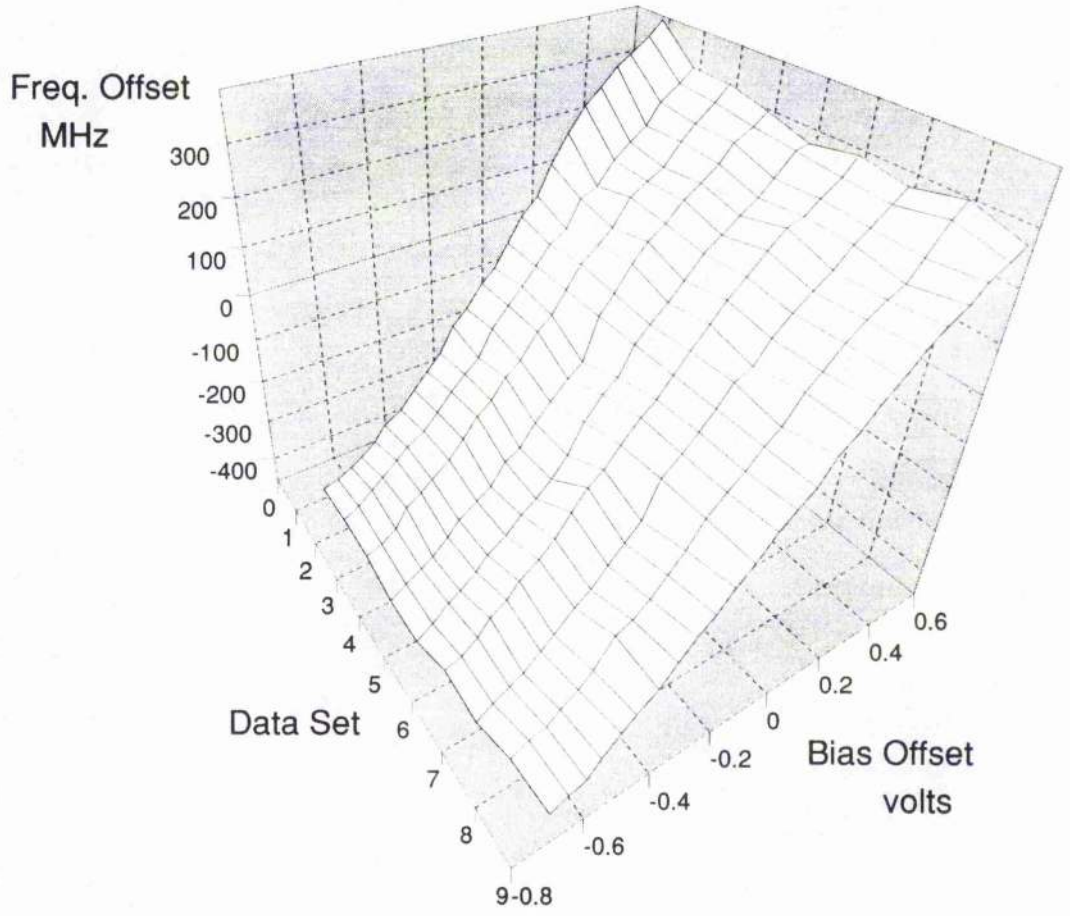


Fig. 4·6 : Frequency Tuning versus Backshort Position

From Fig. 4·6 one can see that the voltage-frequency tuning behaviour is essentially unaffected by the position of the backshort - the shapes of the data sets are fairly uniform. This would be expected since the backshort adjusts the termination of the second harmonic whilst the principal frequency tuning is done at the fundamental which is controlled by the cavity height.

The power versus bias tuning curves are shown in the next

graph, again using the values at 5.0 volts as a reference. The backshort settings are the same as for the previous graph.

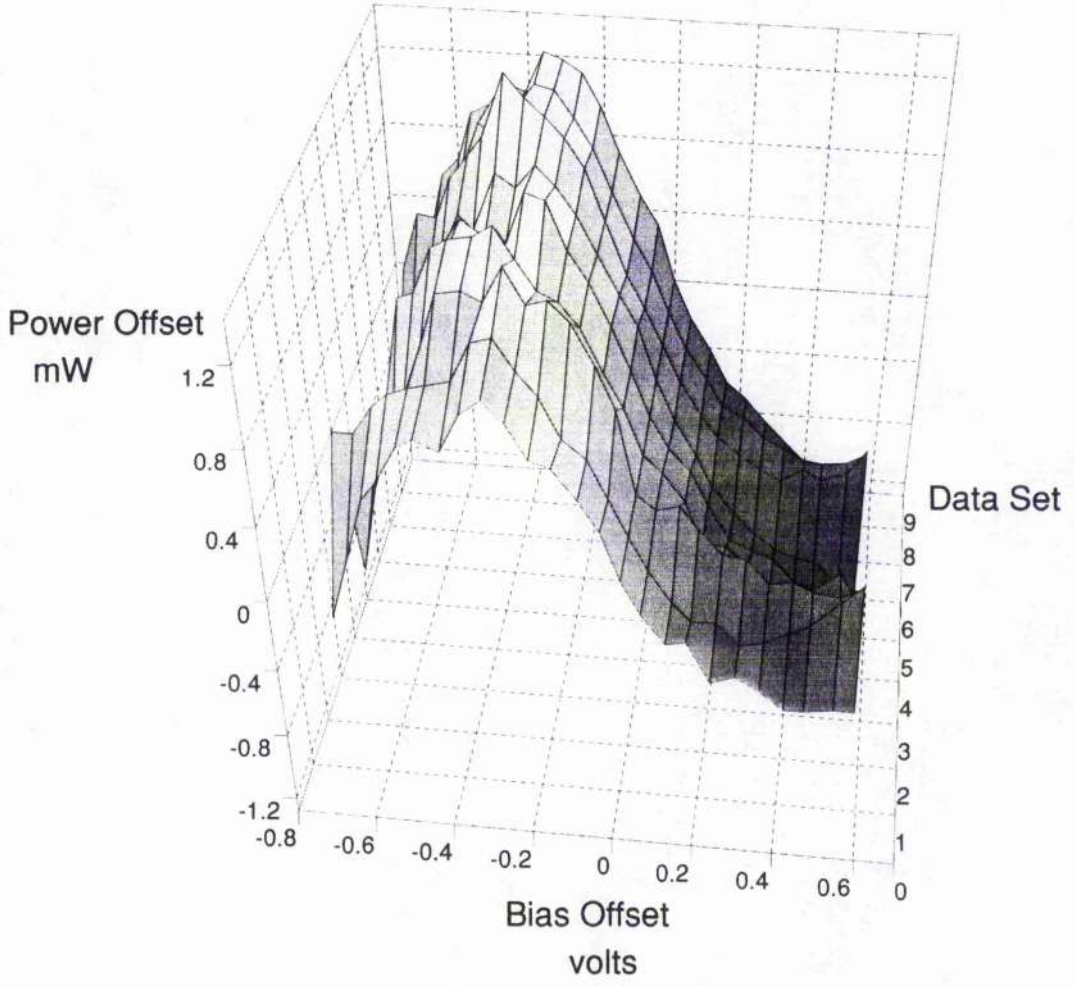


Fig. 4·7 : Power Variation versus Backshort Position

Here again one can see that the shape of the voltage-power tuning curves remain fairly constant with backshort position, though of course the actual power level is controlled with this adjustment.

I am uncertain as to the cause of the fluctuations within individual data sets. Usually, the power curves are quite smooth - compare with Fig. 4·9 for example. Ideally, given sufficient time, these measurements could be repeated and it may be possible to explain this effect.

4.3.3 Choke Varied - Backshort Fixed

A further set of tuning curves was recorded for a range of choke positions, whilst keeping the backshort fixed in position. The voltage-frequency curves are shown below, using 5.0 volts as the reference bias point. The choke positions were 0.80mm, 0.85, 0.90, ..., 1.55, 1.60, 1.70, 1.75, 1.85, 2.00 & 2.05mm. The settings do not increment equally because certain cavity lengths gave rise to bias oscillations at low voltages, so frequency and power measurements could not be made.

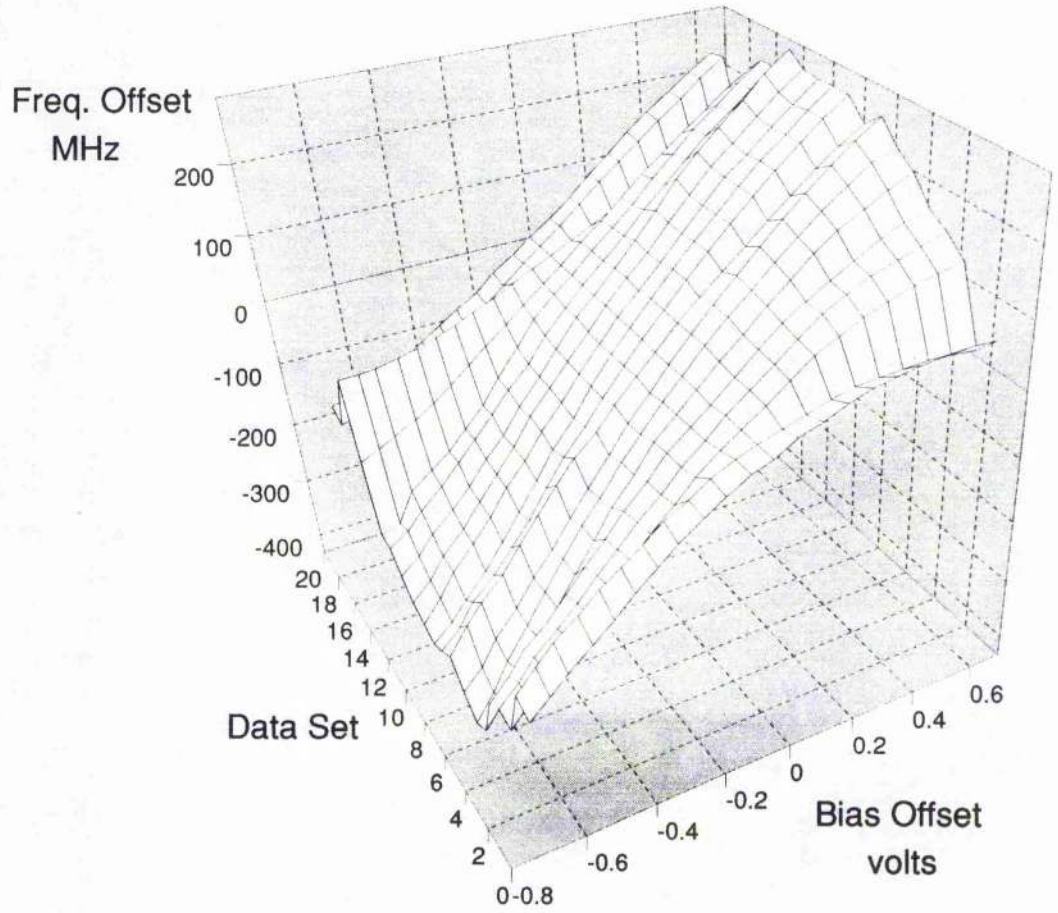


Fig. 4·8 : Frequency Tuning versus Choke Position

This graph shows that there is little in the way of a trend in tuning behaviour other than a slight reduction in frequency sweep at the extremes of the choke positions. The curves in the middle are

most linear and cover the largest frequency range. Note that there is more variation in the shape of the curves in this case than that for the backshort-only case above. This is not unsurprising since adjustment of the cavity length alters the impedance seen by the fundamental frequency of the diode and is thus used as the principal method of frequency tuning. This having been said, there is still little evidence for a general trend. The voltage-power curves for the same data sets are presented below.

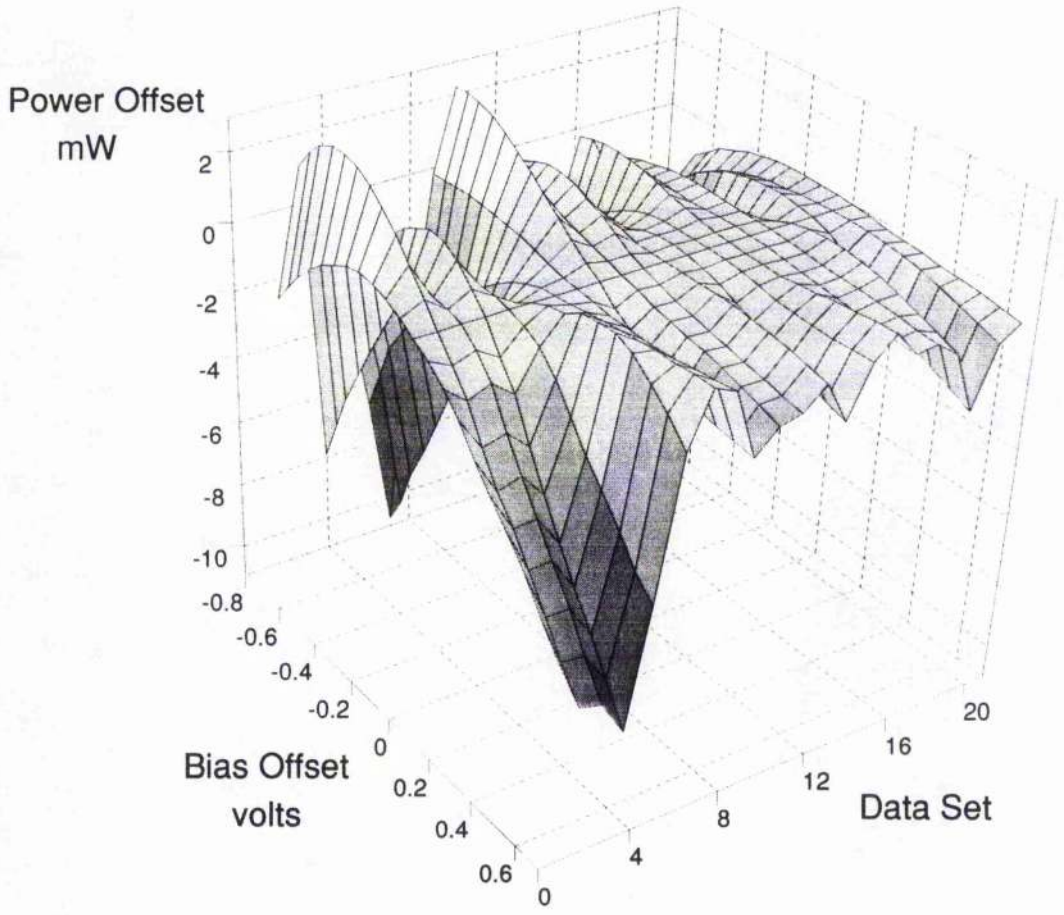


Fig. 4·9 : Power Variation versus Choke Position

Here the curves are particularly disparate showing virtually no macroscopic behaviour. The power tuning seems to fluctuate most at short cavity lengths (low data set numbers) which may be a result of the rapidly changing circuit impedance.

The measurements shown above present very little evidence to suggest that the variation of either the cavity height or the backshort position systematically affect the bias tuning characteristics of this oscillator.

The data recorded for backshort adjustment show there is almost no effect on the frequency tuning of the diode and only a small variation in the general shape of the power tuning curves. The backshort controls the power coupled out of the oscillator, without significant frequency pulling, therefore, one would not expect much influence on the frequency tuning.

In the case of varied choke position, the graphs show more variation from data set to data set, but there are no obvious general trends. The choke position principally alters the resonant frequency of the oscillator cavity but the interaction between the device and the cavity is very complex and it becomes difficult to form a satisfactory explanation.

I would conclude by suggesting that the bias tuning characteristics of an oscillator, particularly the voltage-frequency behaviour, are predominantly a feature of the Gunn device itself rather than its embedding circuit. This view is supported qualitatively through experience of a number of different diode types. The tuning characteristics usually depend on the type of diode rather than on the block in which it is mounted[4]. The tunability of any particular diode will be a function of its nonlinearity which itself results from the device construction. Developing more widely tunable diodes is a task for the device physicist and consequently falls outside the scope of this thesis.

4.4 Dynamic Bias Tuning

Measurements of bias tuning characteristics are traditionally done in a quasi-static fashion whether performed by hand or with an automated system. If an oscillator is being used as a modulator which

is being rapidly voltage-tuned, then one might expect that the tuning properties will be different.

Dynamic bias tuning measurements can be made by linearly ramping the oscillator bias and measuring the power and frequency variations with a calibrated (and preferably linear) discriminator. The operation of such a discriminator is discussed in detail in Chapter 8.

One would start by making quasi-static scans with the Auto_Bias system, gradually decreasing the scan duration and then transferring to the linear ramp and discriminator arrangement to continue measurements up to megahertz modulation rates. The modulation bandwidth is mainly set by the power supply - see Chapter 3.

I began characterising an oscillator in this way using the automated system but found there was no apparent change in the tuning curves. In a scan of 25 points covering 4.3 to 5.7 volts bias, I reduced the scan time from 480 seconds to 90 seconds by reducing the time interval between data points. An example trace of five scans covering this range is shown below. The same oscillator block and diode were used as before.

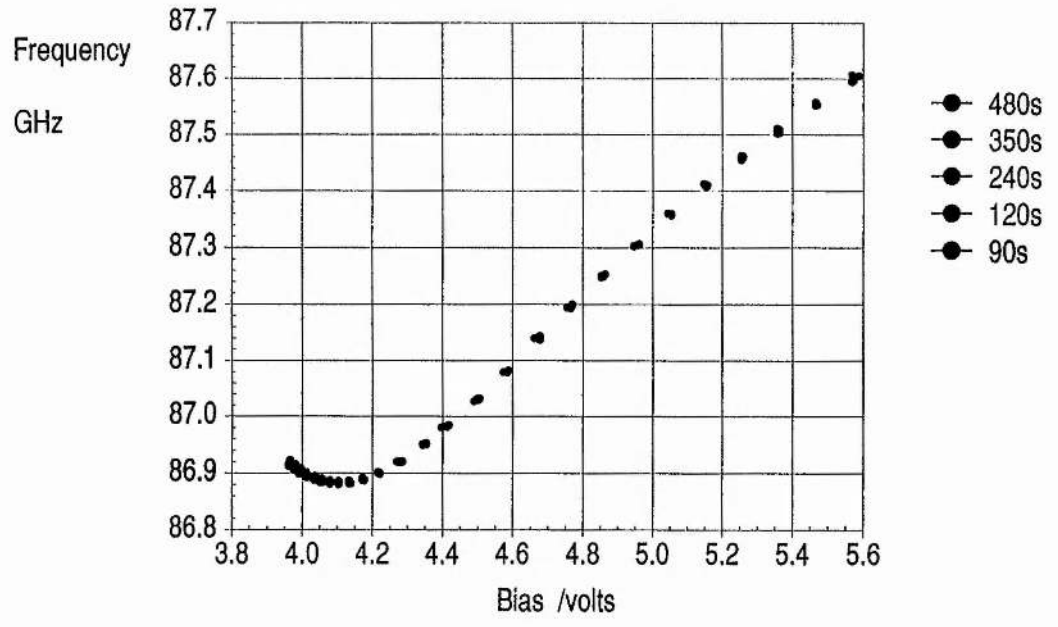


Fig. 4·10 : Consistent quasi-static scans

As can be seen, the points are virtually coincident for all five scans. Therefore, at these timescales there does not appear to be any variation in tuning behaviour with modulation rate. Regrettably, time did not permit any further work in this area and I was unable to continue work on dynamic measurements.

However, this is a topic which deserves attention. On the basis of some extremely brief investigations, it is expected that tuning hysteresis will be evident for an oscillator whose bias is ramped up and down. At slow modulation rates, the hysteresis is large and probably due to thermal effects. As electrical power is cyclically applied, the device temperature varies. Thermal time constants between the device and its package and between the package and the oscillator block then control the output response. At high modulation rates thermal effects will be too slow to respond but a small amount of hysteresis is likely, due to much shorter electronic time constants within the device.

A number of studies of the bias tuning mechanisms of low frequency Gunn diodes have been made, mostly around 10GHz. They concur that at low modulation frequencies (up to about 1MHz for an X-band device) the bias tuning is thermally controlled whereas at higher frequencies the device capacitance becomes the dominant tuning element[5,6,7,8]. The frequency dependence of the modulation sensitivity disappears above about 1MHz, and above 10MHz a dominant phase modulation is observed which is flat up to the relaxation frequency of the rf energy stored in the cavity. This is typically 0.5-1GHz for a coaxial X-band cavity.

DeSa and Hobson have shown[9] from quasi-static measurements that the incremental frequency:voltage change df/dV is proportional to the incremental frequency:temperature change df/dT. At higher modulation rates, up to about 1MHz, they noted two thermal effects. Firstly, a fast device-dependent effect which is a result of the equilibrium between device dissipation and heat flow out of diode and secondly a slower effect due to heat flow from device to heat sink.

Generally speaking, the ac thermal response is largely dictated by thermal response of diode substrate. The use of an integral heat sink (IHS) in diode manufacture allows the semiconductor substrate to be thinned substantially, thereby lowering the thermal impedance and increasing the power handling capability[10].

I would like to investigate this subject in detail in the future with respect to our W-band sources as it will shed light on the operating conditions of modulated oscillators. Undoubtedly, the oscillators in use at St. Andrews differ substantially in operation from the ones described in the references above. X-band diodes are usually transit-time devices mounted in fundamental mode cavities, whereas our oscillators extract the second harmonic frequency of very short devices operating in complicated hybrid modes. It may be possible to model the dynamic bias tuning processes and potentially explain the mechanisms involved. Ultimately, an understanding of this property may allow one to compensate for nonlinearities and produce Gunn oscillator modulators which have more ideal characteristics.

REFERENCES

[1] Intelligent Interfaces, P.O. Box 80, Eastleigh, Hants SO5 5YX.
[2] Wild Vision Ltd., 15 Witney Way, Boldon Business Park, Boldon Colliery, Tyne & Wear NE35 9PE.
[3] DA1349-4 Harmonic Mixer, GEC Plessey Semiconductors, Crompton Road, Groundwell Industrial Estate, Swindon, Wiltshire SN2 5AF.
[4] Dr. Graham Smith, St. Andrews University, Private Communication.
[5] Faulkner, E.A. & Meade, M.L., "Frequency Modulation Sensitivity of Gunn Oscillators", *Electronics Letters*, **5** (10), 15th May 1969, pp. 217-218.
[6] Martin, B. & Hobson, G.S., "High-Speed Phase and Amplitude Modulation of Gunn Oscillators", *Electronics Letters*, **6** (8), 16th April 1970, pp. 244-246.
[7] Martin, B. & Hobson, G.S., "Angle Modulation of Frequency-Locked Gunn Oscillators", *Electronics Letters*, **7** (14), 15th July 1971, pp. 399-401.
[8] Hobson, G.S., *The Gunn Effect*, Oxford, 1974, pp. 99-104.
[9] DeSa, B.A.E. & Hobson, G.S., "Thermal Effects in the Bias Circuit Frequency Modulation of Gunn Oscillators", *IEEE Trans. Electron Devices*, **ED-18** (8), Aug. 1971, pp. 557-562.
[10] Eddison, I.G., "Indium Phosphide and Gallium Arsenide Transferred-Electron Devices", *Infrared & MM Waves*, **11** (Ch. 1), Ed. K. Button, Academic Press, 1984, pp. 1-59.

Chapter Five
Semi-chaotic pulse effects in self-modulated Gunn oscillators

by D.A.Robertson†, G.M.Smith†, J.C.G.Lesurf†, N.R.Couch‡, M.J.Kearney‡

† Dept. Physics, St.Andrews University, North Haugh, St. Andrews, Fife, KY16 9SS.

‡ GEC-Marconi Ltd., Hirst Research Centre, Elstree Way, Borehamwood, Middx, WD6 1RX.

This chapter contains a verbatim copy of an article written by myself and the listed co-authors which was published in the International Journal of Infrared and Millimeter Waves, volume 15, number 3, pages 493 - 503, 1994.

5.1 Abstract

Self-modulation behaviour has been observed in GaAs Gunn oscillators operating at W-band, leading to nanosecond pulses of relatively high peak power. Pulse repetition rates and pulse-widths can be varied reproducibly as a function of bias voltage and bias line termination, and there is evidence for semi-chaotic behaviour in certain regimes.

5.2 Introduction

Gunn diodes operating near threshold often exhibit so-called "bias oscillations" - low frequency oscillations (typically a few megahertz) on the bias line which are usually considered as a form of bias instability, produced because the device negative differential resistance is poorly matched to the bias line impedance. In general,

these oscillations are considered to be undesirable and the diode is operated at higher bias levels to ensure cw operation. Additionally, RC networks and ferrite beads are often used on the bias line, as close to the diode as possible, to help suppress the bias oscillations.

Jaskolskii and Ishii have reported simultaneous low frequency relaxation and high frequency microwave oscillations in a GaAs Gunn diode[1]. The low frequency relaxation oscillations modulated the microwave signal which appeared as a pulsed output. These oscillations were shown to be voltage-dependent and circuit-independent. Low frequency negative resistance in Gunn diodes has been exploited to produce u.h.f. oscillators[2,3] and parametric amplifiers[4]. Similar modes of oscillation have been reported in tunnel diodes.

Similarly, IMPATT diodes are well known to be very susceptible to bias oscillations and bias line noise which frequently leads to burnout of the device. Brackett[5] has shown that the unstable behaviour in IMPATTs is caused by a low-frequency negative resistance which extends from dc to several hundred megahertz. Whilst the negative resistance is unavoidably caused by the avalanche process in the diode, the power supply impedance and bias line termination play a great part in whether or not the circuit oscillates at low frequency. The mechanisms causing low frequency negative resistance in Gunn devices are certainly different from those in IMPATTs but in both cases the bias circuit is a critical factor in determining the behaviour of the oscillator.

Here we are reporting the results of measurements made on novel Gunn devices operating at 80-95GHz, designed and fabricated by GEC-Marconi Ltd., Hirst Research Centre, which have exhibited an enhanced tendency to such effects. These devices are similar to the graded gap injector Gunn diodes reported elsewhere[6,7]. The bias oscillations can correspond to pulsed output of the Gunn diode and can occur for a range of voltages well above threshold. The effects reported here show large amplitude relaxation oscillations with nanosecond rise and fall times of high peak power, and semi-chaotic behaviour is evident under certain conditions.

5.3 Apparatus

The Gunn oscillator was of a tunable second harmonic design after Carlström *et al*[8], and Smith[9]. A quasi-optical arrangement was favoured for observing the pulsed behaviour to ensure electrical isolation between source and Schottky-barrier detector. Additionally, both a waveguide isolator, mounted between the oscillator block and feedhorn, and a free-space ferrite isolator were utilised to ensure the oscillator would not be loaded by any signal reflected from the detector. Fig. 5·1 shows the experimental layout.

The detector utilised was a harmonic mixer designed for an EIP microwave counter (GEC Plessey Semiconductors model DA1349-4). Whilst not specifically designed as a power detector it was adequately sensitive and had a sufficiently wide bandwidth. The detector was biassed with a battery supply via a Picosecond Pulse Labs 5575A bias tee to ensure maximum sensitivity.

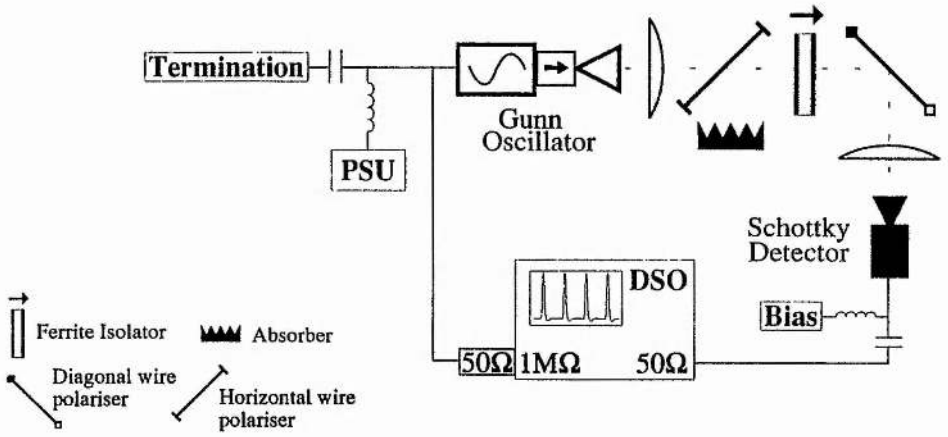


Fig. 5·1 : Apparatus

The detector output signal was then measured on both a fast digital storage oscilloscope (DSO) and a number of spectrum analysers which were available. The DSO employed was a four gigasample per second HP54720A which gives a real bandwidth of about 1GHz - ideally suited for observing nanosecond pulses. Spectral measurements of the pulse signals were made with either an HP8590A spectrum analyser

(1.5GHz bandwidth) or an HP70000 series analyser which had a full bandwidth of 2.9GHz.

Bias for the Gunn diode was supplied from one of our in-house designed power supplies which had a modulation bandwidth of 100kHz and a low frequency output impedance of ~0.65Ω for bias voltages above threshold. The bias was fed to the oscillator block through a high-current bias tee which served to isolate the power supply from the diode and permitted adjustable termination of the bias line on the r.f. arm of the tee. Using the bias tee showed that different types of power supply made little difference on the behaviour of the oscillator. The bias line into the oscillator was also monitored on the DSO, with a 50Ω pad being used to match the line.

5.4 Pulse Effects

The oscillator was found to produce a wide variety of pulse effects depending on a number of parameters. Pulse shapes were a function of the termination of the fundamental and second harmonic frequencies of the diode, the bias voltage, and the termination of the bias line. The fundamental frequency of the oscillator was tunable by altering the height of the resonant cavity above the diode, whilst the second harmonic was tuned by altering the position of a sliding waveguide backshort.

Tuning the fundamental tended to produce either short pulses (3-4ns, see Fig. 5·2), broad pulses (200-800ns, see Fig. 5·3) or broad pulses with a sharp spike on the rising edge. The pulses jumped suddenly from one type to another as the cavity length was adjusted but the pattern of changes was quite repeatable with cavity length (a small hysteresis was noted depending on whether the cavity was tuned up or down). Adjusting the second harmonic usually just altered the amplitude of the pulses (i.e. varied the power coupled out of the cavity) but sometimes tuned one pulse shape into another, indicating a small

amount of frequency pulling. The discontinuous changes between pulse shapes when tuning the fundamental are probably due to the large frequency change with changing cavity length. Additionally, the mechanical tuning of the cavity is not totally smooth.

Pulsed output occurred for mean bias voltages between about 3 and 5 volts and over this range the pulse repetition period decreased with increasing bias, typically from 800 to 150ns. At higher bias levels the diode output went into cw operation. The shortest pulses observed were of the order of 3.5ns long with a 1.2ns risetime and these values stayed roughly constant with voltage. As the bias was increased from below threshold, single widely spaced pulses appeared, moved closer together and then transformed into bunches of two, three or more pulses separated by longer intervals. The bunches then formed into a continuous pulse train up until voltages just below the onset of cw operation. At this point the diode produced bursts of pulses separated by relatively long gaps, and these gaps were voltage dependent. These bursts looked typical of a relaxation oscillator (see Fig. 5·4).

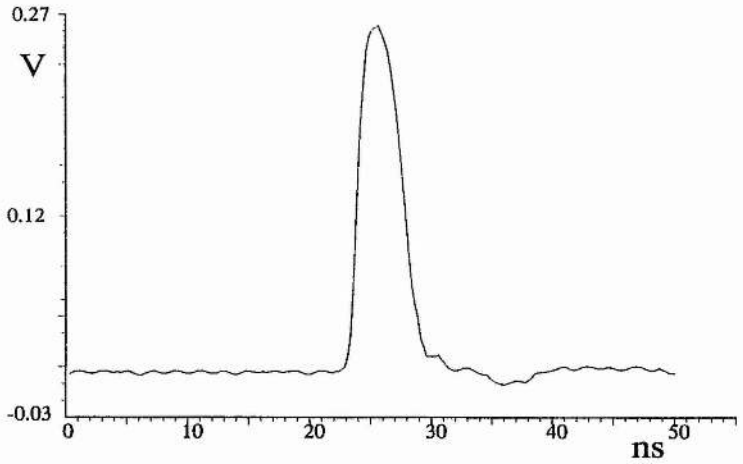


Fig. 5·2 : Typical 3.5ns pulse

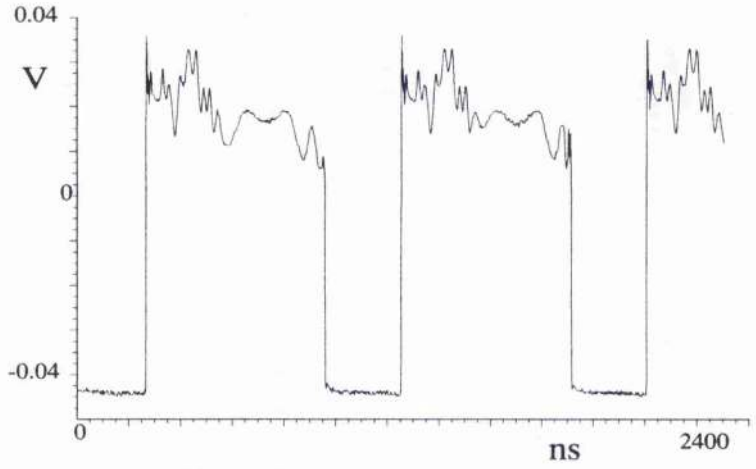


Fig. 5·3 : An example of broad pulses

Just at the transition to cw output, the oscillator showed a low-amplitude sinusoidal oscillation of around 750MHz which gave rise to classic frequency modulation of the signal. This is discussed in more detail in the section on heterodyne measurements.

A degree of hysteresis was noted when tuning the voltage up and down. The pulses appeared and disappeared at slightly lower voltages when the bias was decreasing. A greater difference was noted at the low-bias threshold: when the bias was decreased, pulses were observed for voltages below the level at which the diode turned on when bias was increased.

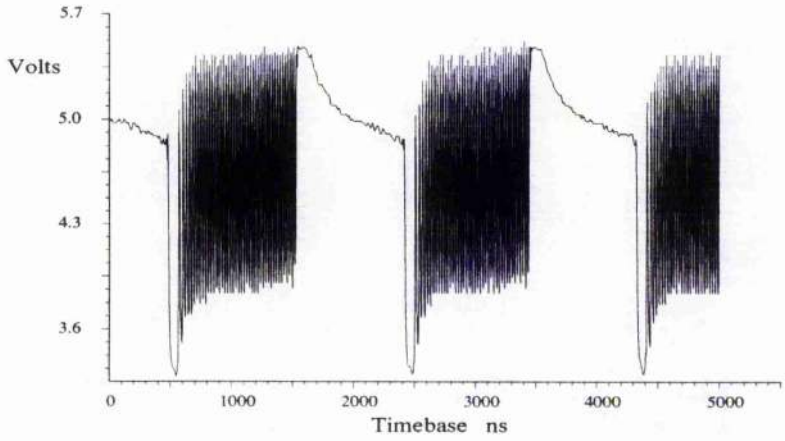


Fig. 5·4 : Bursts of oscillation on the bias line

In conjunction with time domain measurements, the pulse signal

was measured in the frequency domain with a spectrum analyser. The spectra showed all the hallmarks of typical pulsed waveforms: a comb of harmonics spaced at the pulse repetition frequency, shaped by a lobed envelope, with the lobe width corresponding to the inverse of the pulse width (see Fig. 5·5). Double pulsed behaviour showed two sets of harmonics corresponding to the two different (inter-pulse and inter-bunch) period values.

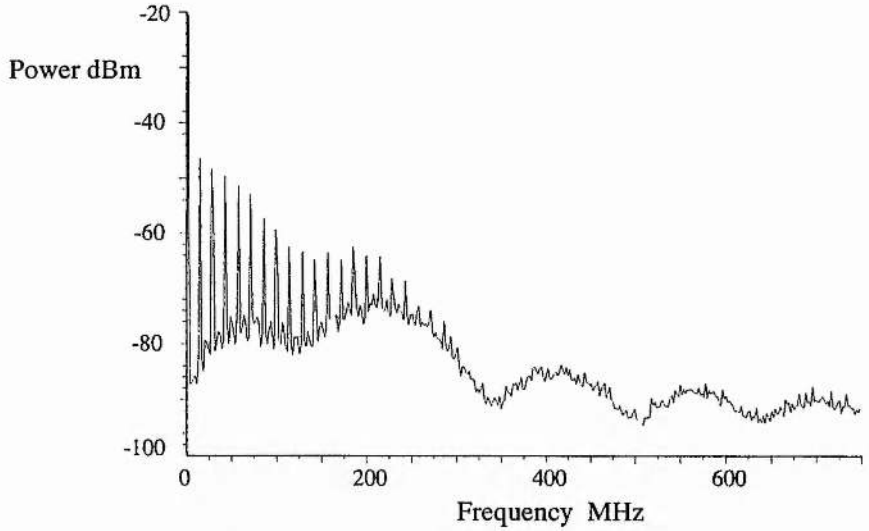


Fig. 5·5 : Pulse spectrum

Monitoring the Gunn diode bias line showed very large voltage excursions corresponding to the pulsed output. These were typically of the order of 2.0 to 2.5 volts in magnitude. The bias line pulse corresponding to the short pulse illustrated above had a risetime and pulsewidth of 3.2 and 9.0ns respectively. In general, the bias line pulses were slower and longer than the output pulses. In one case, the bias voltage swung over 1.8V peak-to-peak with 500ns period and a duty cycle of ~50% whilst the output pulses were 3.2ns wide with 1.1ns risetime and only occurred on the rising edges of the bias pulses (i.e. at 500ns period). The bias pulses were swinging over a sufficiently large range that the diode was alternately turning almost fully on (nearly cw) then fully off, and the sharp voltage increase gave rise to a pulse which modulated the microwave signal.

The bias line pulses changed shape with adjustment of the fundamental frequency and the inter pulse period was voltage dependent, in similarity with the output pulses. Multiple pulse bunches, bursts of oscillations and small amplitude oscillations were all observed on the bias line corresponding to such behaviour as seen by the detector.

Pulse behaviour was dependent on the matching circuit mounted in the oscillator block. This normally consists of a zener diode, to prevent over-biassing, in parallel with a resistor and capacitor to ground and quench oscillations. Typical values of the components are V_Z = 6.8V, R = 22Ω and C = 0.01μF. The zener tended to limit the amplitude of the bias line pulses but it sometimes gave pronounced ringing on the pulse edges. Henceforth, the zener was removed from the oscillator to simplify the circuit. The RC snubber was retained, however, to try to match the 50Ω cable used to supply the bias.

Different bias tee terminations were found to have marked effect on the pulse behaviour. Connecting unterminated lengths of cable to the r.f. arm of the tee produced bunches of pulses, with more pulses per bunch for longer lengths of cable. It was possible to generate 17 pulses or more in any one bunch where there was only one pulse for a terminated cable. Using a termination restored the single pulses and made the pulse behaviour independent of the length of the cable. Clearly, the voltage pulses at the diode were being reflected from the open circuit end of the line and returning to the diode where they generated further output pulses. Normally, the r.f. arm of the tee was terminated in a 50Ω load when observing pulses.

5.5 Power Measurements

Power measurements on the oscillator were made with both a thermal detector (Anritsu ML83A) and a fast diode detector (Boonton 4220). Additionally, the detector diode used for monitoring pulses was

calibrated with a cw source against the Anritsu, driving into a 50Ω load. The calibration was done with the detector unbiassed, for which the Schottky diode has a threshold of about 1.8mW. CW calibration was not performed with the detector biassed as the bias tee output is ac coupled.

Just above threshold the peak power could exceed as much as 50mW whilst the average power was less than 0.5mW. In general, the peak power was lower than this but remained fairly constant with bias whilst the average power increased slowly.

The graph below (Fig. 5·6) shows results taken with all three detectors. The Anritsu measures average (thermal) power and shows clearly the gradual increase in power from turn on to cw operation. The slight discontinuity around 5.2V is not fully understood but it may be that the lower section (3.2 to 5.2V) under-reads because the pulses were too fast for the bolometer sensor which has a response time of about two seconds. The Boonton measures peak power but the values here don't include a calibration factor of about 1dB. With this in mind there is good agreement with the peak power measured with the diode detector - for this, the measured peak voltage was converted to a power with a calibration curve. The calculated average power was obtained by measuring the average voltage of the pulse train and then converting to milliwatts. The peak and average voltages were calculated with the measurement functions of the HP54720A scope. Note that here the detector was unbiassed and had a threshold level of about 1.8mW so the bottoms of the pulses were not resolved and this probably accounts for the high level of the 3.2 to 5.2V section. Note that the peak, calculated average and Anritsu measurements all agree closely over the cw region.

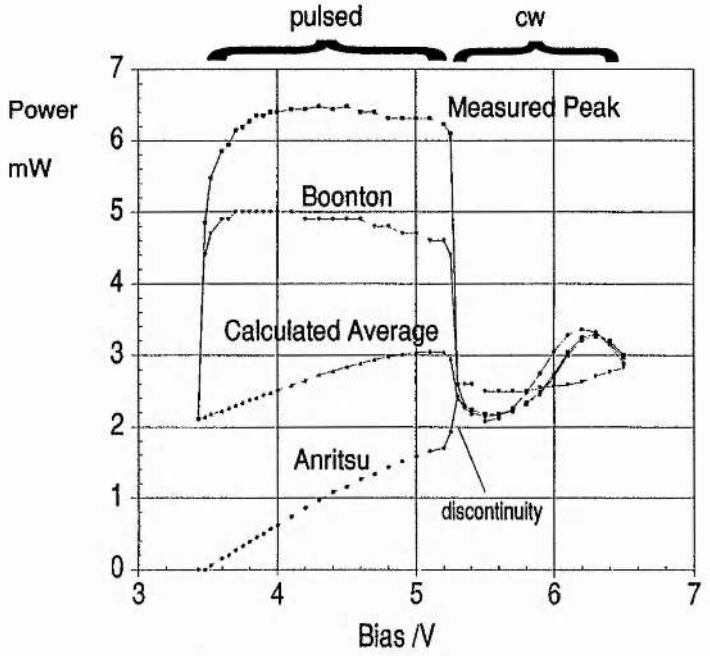


Fig. 5·6 : Power Measurements

These measurements were taken with the oscillator connected to the power detectors on the straight arm of a 10dB coupler. Thus, there is a loss of perhaps as much as 3dB which has not been included in the values here.

5.6 Heterodyne Measurements

Interferometric measurements were performed on the Gunn oscillator with a Martin-Puplett Interferometer and it was hoped that such measurements would show any signal modulation or wide band effects due to the pulsed output. However, the Fast Fourier Transforms obtained did not show any such details. The resolution of the FFT was 1.4GHz so any features narrower than a few gigahertz would not appear.

In order to investigate the spectral details close to the signal frequency, heterodyne measurements were made on the oscillator by mixing the output with a cw local oscillator in a Schottky barrier diode.

The pulsed oscillator output was coupled into a free-space

Gaussian beam and passed through a ferrite isolator. The beam was then coupled into a receiving feedhorn mounted on the -10dB arm of a 10dB coupler. The local oscillator was mounted to the direct arm of the coupler through a waveguide isolator to ensure minimal frequency pulling of the IF by the pulse signal. The diode detector (the same unit described above) was mounted at the other end of the coupler, biassed through a bias tee and the output fed to an HP70000 series spectrum analyser with 2.9GHz full bandwidth.

With the oscillator giving cw output, the IF peak was clearly visible, about 50dB above the noise floor. Reducing the bias voltage caused small sidebands to appear around the peak. These corresponded to a small amplitude sinusoidal oscillation (typically around 90MHz) observed on the bias line. The oscillation amplitude increased with decreasing voltage giving rise to as many as 16 sidebands extending over 1300MHz, each spaced at the pulse repetition frequency, in a clear example of narrow band frequency modulation (Fig. 5·7).

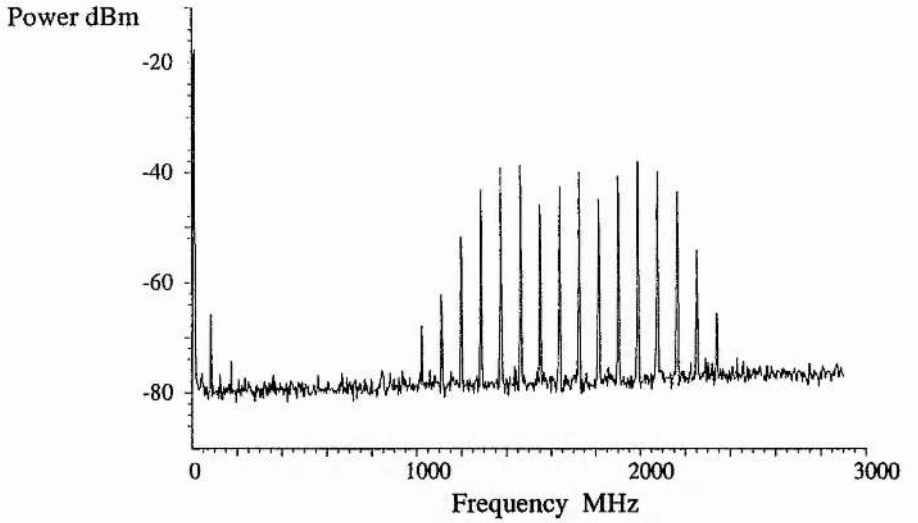


Fig. 5·7 : IF sidebands

Lowering the bias further brought the diode into bursts of oscillations, but with the small amplitude sinusoidal signal superimposed on top. The spectrum then became very wideband, covering nearly 3GHz, with a rather indistinct "humped" shape

extending to one side of the sidebands. These sidebands disappeared at lower bias levels and gave way to a plateau shaped bump located where the IF peak had been, reminiscent of wideband sinusoidal FM. This hump was evident as long as the oscillator was producing bursts of oscillations and disappeared when the output became purely pulsed. The IF spectrum of pulsed output consisted of a series of approximately a dozen humps, each increasing in amplitude from one end to another and each with a width of about 200MHz. Below is shown a spectrum of output for bursts of oscillations indicating all the main effects (Fig. 5·8).

Note the residual IF peak and pair of sidebands, the "crown" shape characteristic of wideband sinusoidal FM and the series of humps extending to almost 3GHz.

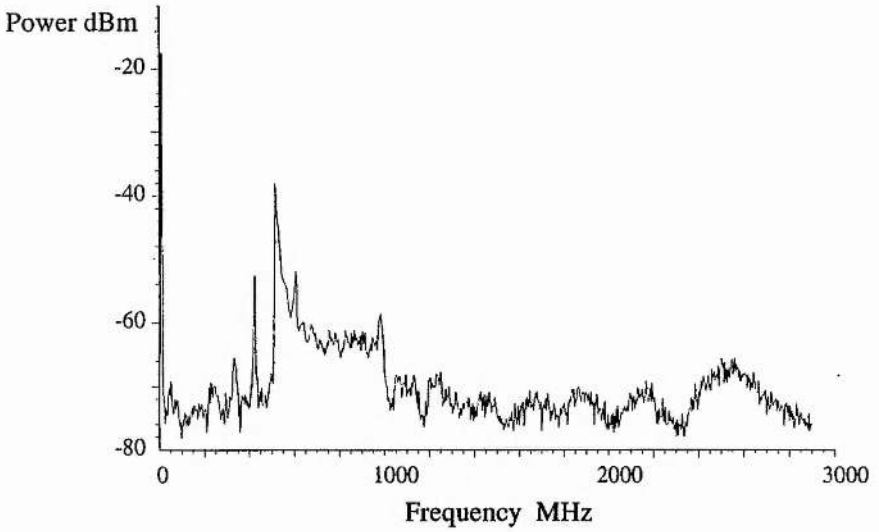


Fig. 5·8 : IF spectrum of pulsed oscillator

The oscillator was also set to produce broad pulses (~450ns wide) in order to examine the associated frequency sweep. The spectrum obtained was similar to that shown above but the crown shape was 1000MHz wide and the humps extended to well over 3GHz, beyond the span of the spectrum analyser. Clearly, there is an extremely large frequency spread associated with the broad pulses.

Reducing the span and resolution bandwidth when examining

the wide spectra did not show any clear frequency components suggesting that the signal is fairly noisy in nature.

5.7 Semi-Chaotic Effects

The multi-pulse bunches observed at certain bias levels indicated the possibility of semi-chaotic behaviour in the Gunn diode. In order to examine this effect carefully, digitized scope traces of the pulse signal for a range of bias voltages were saved to disc and loaded into a computer. Software was written to calculate the inter-pulse spacings for each trace and plot them against bias voltage. The resulting plot is a bifurcation diagram which shows periodicity of the pulse waveform against bias (Fig. 5·9).

As can be seen below, the period is large at low voltages and decreases as the bias increases to 3.95V. Above that, the oscillator gave bursts of oscillation which were not considered as multi-periodic waveforms. At each end of the plot, the data points are essentially single valued but some are spread slightly due to jitter in the pulse waveform. At about 3.35V there is a very clear bifurcation from period-1 to period-2 behaviour. This period-2 feature is evident up until around 3.49V when single period behaviour suddenly returns. Within this region there is evidence for period-3 and possibly period-4 behaviour but the effect is blurred by pulse timing jitter. Note that this diagram shows all the measured inter-pulse spacings in each digitized trace and does not indicate the number of pulses per bunch.

Whilst these results are a little unclear in detail, they do show the general pattern of voltage-dependent behaviour which this particular Gunn device exhibits. The multi-period waveforms observed are a strong indication of semi-chaotic behaviour in the Gunn oscillator circuit.

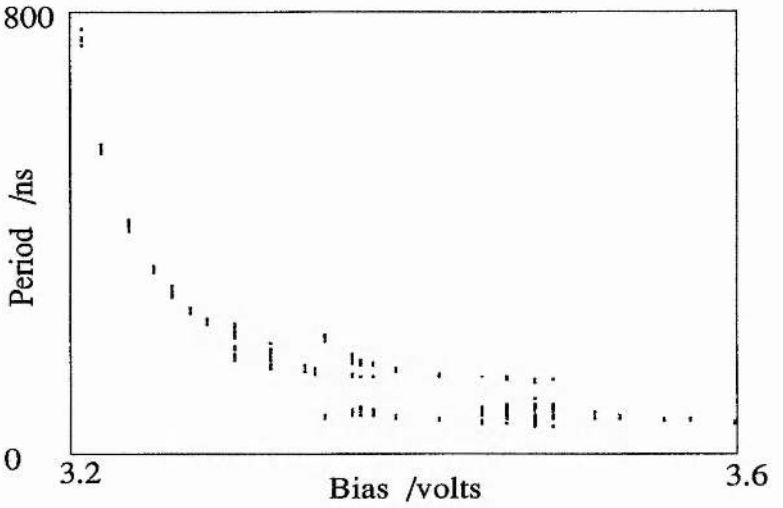


Fig. 5·9 : Bifurcation Diagram

Numerical models of Gunn diodes by Lesurf & Robertson have shown similar behaviour to that shown above[10]. These models were based around three essential features for simulating Gunn effect devices: a nonlinear IV function, a voltage-dependent capacitance and a response delay in the device conductance. The results of the simulations exhibited period doubling bifurcations, multi-period waveforms and chaotic output. Relaxation type oscillations and burst effects were also generated by the models. In parallel to these computations, simple analog models of Gunn diodes were constructed which also showed semi-chaotic behaviour. Mosekilde *et al* describe a fairly detailed numerical simulation of nonlinear dynamic phenomena in Gunn diodes in which mode locking and spatiotemporal chaos are reported[11]. However, their work modelled lower frequency Gunn diodes (10 to 20GHz) operating in domain based modes whilst the Gunn diodes measured here operate at much higher frequencies (90 to 100GHz) possibly in a hybrid field-accumulation mode.

5.8 Discussion

We suggest that the mechanism involved in producing these pulses is similar to that described in Ref. 1. Whilst the diode is

oscillating at W-band there is also significant negative resistance at lower frequencies. When suitably biassed, the diode undergoes relaxation oscillations at low frequency, swinging in and out of the r.f. negative resistance region thus modulating the millimetre wave output.

The low frequency negative resistance is a property of the diode in its embedding circuit and is strongly voltage dependent. The pulse repetition rate varies with bias. The pulse shapes are determined by the properties of the surrounding millimetre wave circuit, particularly the terminations of the fundamental and second harmonic frequencies. It has been shown that the biassing circuit plays a crucial role in the behaviour of the pulses indicating a clear interaction between the diode negative resistance and bias line impedance. The pulse effects are particularly prominent in diodes which exhibit a strong frequency-bias dependence. This suggests that diodes with steeply nonlinear IV and CV characteristics are likely to operate in multi-periodic or semi-chaotic regimes.

The fast pulses are not subject to the thermal limitations of cw operation so the peak power is relatively high (typical values are 50mW peak as opposed to 30mW cw for diodes of this type).

Further work is needed to understand this self-modulation behaviour in detail, representing as it does a subtle interaction between the diode and its circuit.

5.9 Possible Extensions

In order to reach a greater understanding of the effects reported here, pulsed Gunn diode behaviour needs to be investigated in more detail. For example, the low frequency properties of the diode could be investigated under conditions in which the the r.f. oscillation is quenched - this could be done by introducing a lossy material (e.g. Eccosorb) into the resonant cavity, hence suppressing the r.f. field. This may then allow careful matching of the bias line and optimisation of the

RC network located in the block. As a result, a greater understanding of the operation of the diode and its interaction with the surrounding circuit would be reached. Knowledge of the pulse generating mechanisms would allow any unwanted oscillations to be suppressed where necessary. It may also be possible to lock the pulse train to an external signal if a signal of a frequency close to that of the pulse train is injected into the ac arm of the bias tee. This might enhance the timing stability of the millimetre wave output pulses, which could lead to the construction of a fast-pulse millimetre wave source that would have applications in radar and communications. A self-pulsing oscillator would be attractive since no external modulator is required and, as a source of semi-chaotic pulses, could be utilised as a noise source. Finally, further numerical simulations should also be explored in order to model the complex nonlinear behaviour that Gunn devices are capable of producing.

Acknowledgements

This work was supported in part by the MOD Procurement Executive, sponsored by DRA Malvern and in part by the National Physical Laboratory, Malvern. We would also like to thank all those who loaned equipment used in this work, especially Pete Hirst for allowing extended use of the HP54720A oscilloscope.

REFERENCES

[1] Jaskolski, S.V. & Ishii, T.K.: "Simultaneous low-frequency relaxation and high-frequency microwave oscillation of a bulk GaAs c.w. oscillator", *Electronics Letters*, 1967, **3**, pp. 12-13.

[2] Cawsey, D.: "V.H.F. and U.H.F. Gunn-effect Oscillators", *Electronics Letters*, 1967, **3**, pp. 550-551.

[3] Brunt, G.A.: "Low-frequency negative resistance of X-band Gunn diodes", *Electronics Letters*, 1969, **5**, pp. 151-153.

[4] Aitchison, C.S. *et al*: "Self-pumped Gunn-effect parametric amplifier", *Electronics Letters*, 1969, **5**, pp. 36-37.

[5] Brackett, C.A.: "The Elimination of Tuning-Induced Burnout and Bias-Circuit Oscillations in IMPATT Oscillators'", *B.S.T.J.*, **52** (3), March 1973, pp. 271-306.

[6] Spooner, H. & Couch, N.R.: "Advances in Hot Electron Injector Gunn Diodes", *GEC Journal of Research*, 1989,**7** (1), pp. 34-45.

[7] Couch, N.R. *et al*: "High performance Graded Gap AlGaAs Injector GaAs Gunn Diodes at 94GHz", *IEEE Electron Device Letters*, 1989, **10** (7), pp. 288-290.

[8] Carlström, J.E. *et al*: "A continuously tunable 65-115GHz Gunn Oscillator", **MTT-33** (7), July 1985, pp. 610-619.

[9] Smith, G.M.: "TEOs at MM-Wave Frequencies and their Characterization using Quasi-Optical Techniques", Ph.D. Thesis, St. Andrews, 1990.

[10] Lesurf, J.C.G. & Robertson, M.R.: "Report on the Feasibility of 'Chaotic' MM-Wave Noise Sources", Report to the NPL, 1992.

[11] Mosekilde, E. et al: "Mode locking and spatiotemporal chaos in periodically driven Gunn diodes", *Phys Rev B*, **41** (4), Feb 1990, pp. 2298-2306.

Chapter Six
Spread Spectrum Techniques

As certain regions of the electromagnetic spectrum become increasingly crowded with communications channels, attention moves to lesser used frequencies. Typically, this means a move to higher frequencies and certain communications applications are gradually moving into the millimetre wave region.

The high carrier frequencies of millimetric signals offer a number of advantages to the communications engineer. Apart from the (present) relative freedom from spectral overcrowding, millimetre waves offer wide bandwidths and line of sight operation. Additionally, the short wavelength means that antennas are small and entire systems can remain relatively compact.

Since the late 1940s[1], wideband communications systems have received increasing attention in order to tackle a number of the limitations associated with conventional narrowband systems. Wideband or "spread spectrum" systems as they are known, allow multiple users to share the same frequency space, offer good rejection of unwanted signals and provide message screening or even covert transmission of information.

This chapter introduces the fundamentals of spread spectrum techniques and as such marks the beginning of the final section of this thesis which concludes with the description of a millimetre wave system that combines spread spectrum and quasi-optical signal processing techniques.

6.1 An Introduction to Spread Spectrum

A spread spectrum signal is one which has a bandwidth much greater than is strictly necessary for the transmission of a given quantity of information.

Conventional FM transmission as used in commercial radio broadcasts can be considered as spread spectrum since the rf bandwidth is greater than the minimum necessary for the transmission of audio signals. However, a more rigorous definition of spread spectrum systems specifies that some signal or operation other than the information is used to spread the signal[2]. The techniques described here fit into this category.

Generally speaking, in a spread spectrum system, the information is imposed on a frequency-spread carrier and transmitted, then received, whereupon the frequency spreading is removed and the information is extracted. There is an enhancement in signal-to-noise ratio to be had from this modulation and demodulation process and this is the basis on which such systems are applied. This fact can be elicited from Shannon's equation for channel capacity:

$$C = W \log_2\left(1 + \frac{S}{N}\right) \qquad 6.1$$

where C is channel capacity in bits per second, W is bandwidth in hertz and S and N are signal and noise powers respectively. By expressing eqn. 6.1 in natural logs and using the logarithmic expansion (for small S/N) one arrives at the expression

$$W = \frac{CN}{1.44S} \qquad 6.2$$

From this we can see that the information rate is proportional to the signal bandwidth for a constant signal-to-noise ratio, or alternatively, for a channel with a given capacity, decreasing the bandwidth has the effect of increasing the signal-to-noise ratio.

6.1.1 Process Gain and Jamming Margin [3]

The enhancement in signal-to-noise ratio in a spread spectrum system is termed the process gain, G_P. The despreading of the signal at the receiver is accomplished by correlating the incoming signal with a locally generated frequency-spread reference. When the signals are matched, the desired signal collapses down to its

original pre-spread bandwidth and any uncorrelated noise or interference is spread out to the bandwidth of the local reference. The desired signal is then bandpass filtered and the majority of the noise is rejected. The ratio between the output and input signal-to-noise ratios at the receiver is defined as the process gain:

$$\left(\frac{S}{N}\right)_O = G_P\left(\frac{S}{N}\right)_I \qquad 6.3$$

As a rule of thumb, the process gain is the ratio of the transmission bandwidth to the information bandwidth:

$$G_P = \frac{B_{RF}}{B_{info}} \qquad 6.4$$

The process gain is a property of an ideal receiver and does not take into account the real-life problems of interference, noise and deliberate jamming signals. The ability of a receiver to perform in the presence of a jammer (by which I will take to mean *any* unwanted signals, whether deliberate or otherwise) is expressed as the jamming margin, M_J. This is defined as follows:

$$M_J = G_P - \left[L_{SYS} + \left(\frac{S}{N}\right)_O\right] \qquad 6.5$$

where L_{SYS} is the total receiver loss. For example, a system with a process gain of 25dB, an output signal-to-noise ratio of 10dB and system losses of 2dB would have a jamming margin of 13dB. In other words, it would still be able to function in the presence of jamming signals up to 13dB greater than the desired signal level. Hence, the jamming margin is a realistic measure of how well a system will work in practice.

The limit of operability of a receiver occurs when the jamming threshold is reached. That is defined as the point at which the receiver drops by 1dB from its ideal jam-to-signal (J/S) performance. This is illustrated below.

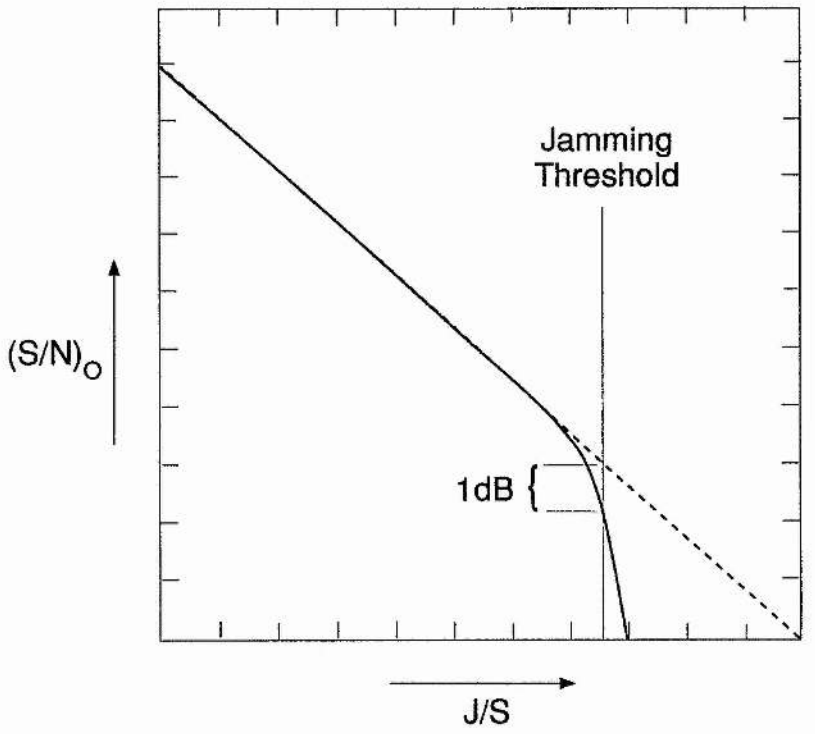


Fig. 6·1 : Typical Process Gain Curve

Through the properties described above, spread spectrum systems offer very flexible means of rejecting unwanted signals. The way in which any particular system is implemented depends on the application in question. The next sections look at different types of spread spectrum systems, their implementation and some of the resulting applications.

6.2 Main Types of Spread Spectrum Systems

6.2.1 Direct Sequence

The best known and most widely used form of spread spectrum techniques is direct sequence (DS) in which a carrier signal is modulated by a code sequence. This can be done by amplitude modulation (AM), frequency modulation (FM) or phase modulation (PM). The most common method is 180° biphase phase-shift keying (BPSK) in which the carrier phase is modulated by a digital pseudo-

random code sequence. One carrier phase represents a "1" and the other phase represents a "0" of the pseudo-noise sequence. Other types include pulse amplitude modulation (PAM), quadriphase phase-shift keying (QPSK) and minimum shift keying (MSK)[4].

A BPSK spread spectrum signal is characterised by a sinc-squared power spectrum with nulls at integer multiples of the code clock frequency and a suppressed carrier - Fig. 6·2. The lack of a carrier means the signal is harder to detect, there is more power available for information transmission and the power envelope remains constant.

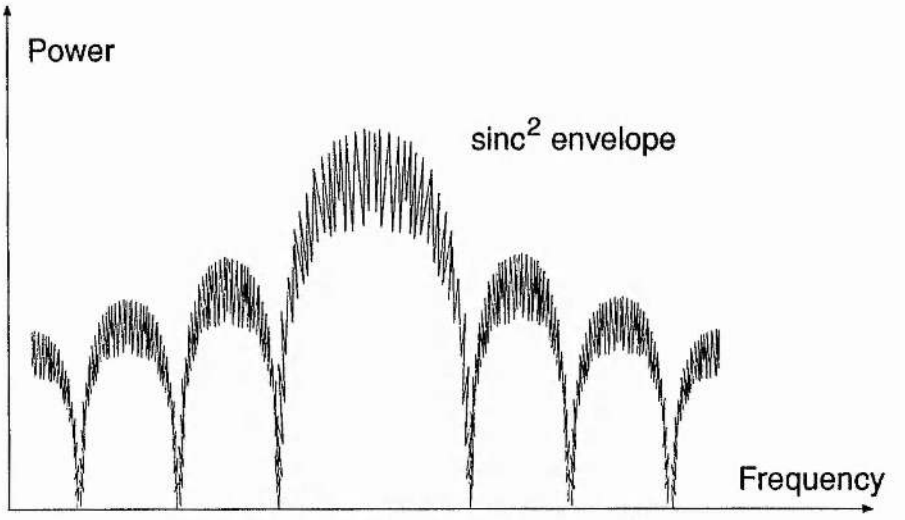


Fig. 6·2 : Direct Sequence BPSK spectrum

In the transmitter, narrowband digital information is combined with the wideband digitally generated code sequence and then used to modulate the carrier. The fast bit rate of the code sequence spreads the carrier over a wide bandwidth. Note that the clock rate of a code generator is called the 'chip' rate, and each bit of the code sequence is called a 'chip'.

At the receiver, the incoming signal is mixed with a locally generated reference consisting of an identical code generator to that found in the transmitter and a frequency-offset local oscillator. When the local reference and signal are out of synchronism, the signal is again spread out to the bandwidth of the local reference.

However, when the two *are* synchronised, the signal collapses down to its original bandwidth and passes through a narrowband IF filter. The IF signal can then be demodulated and the information extracted just as in a conventional radio receiver. Fig. 6·3 schematically represents a complete direct sequence system.

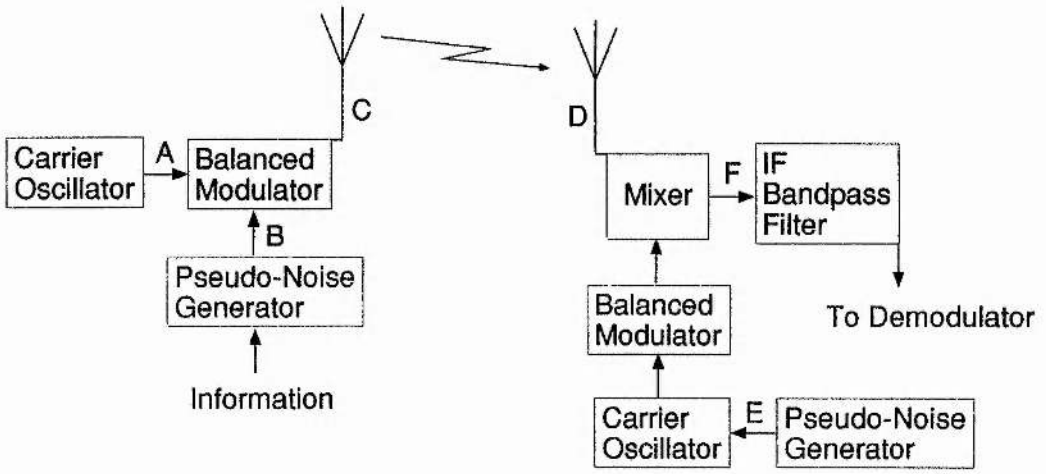


Fig. 6·3 : Direct Sequence System

The waveforms in Fig. 6·4 represent the RF carrier modulation and demodulation process showing the carrier phase inversion at the logic transitions of the PN code. The letters A to F refer to the points in the system as labelled above.

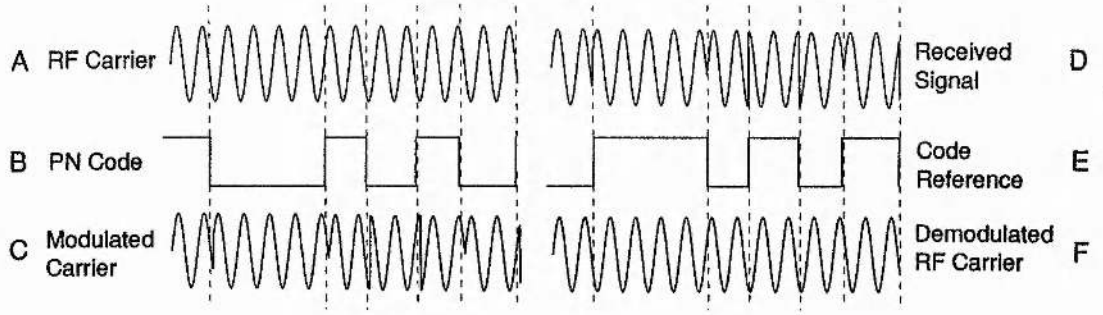


Fig. 6·4 : Direct Sequence Waveforms

The most common method of carrier inversion modulation is to use a double balanced mixer (DBM). The DBM must have a symmetrical response to ensure correct cancellation of the carrier. Any imbalance in the modulation will result in poor carrier

suppression. Additionally, the code sequence must be balanced, in the sense that there should be an equal weighting one 1's and 0's and the logic circuit driving the DBM should have balanced behaviour - e.g. flip-flops should have equal rise and fall times. If code imbalance is present then there will be poor suppression of the code clock and harmonics of the clock will appear at the nulls of the DS spectrum. This makes the spectrum quite obvious to a casual observer.

6.2.2 Frequency Hopping

The second most popular system for spread spectrum communications is known as frequency hopping (FH). It is a form of frequency shift keying (FSK) but with many frequencies rather than the conventional two.

The signal is frequency spread by controlling a frequency synthesiser which hops to any one of a total of N frequencies in a random fashion in response to a pseudo-random code sequence. Ideally, the synthesiser hops instantaneously to a single frequency and the time averaged spectrum is perfectly rectangular with the transmission distributed evenly in every available frequency channel - Fig. 6·5.

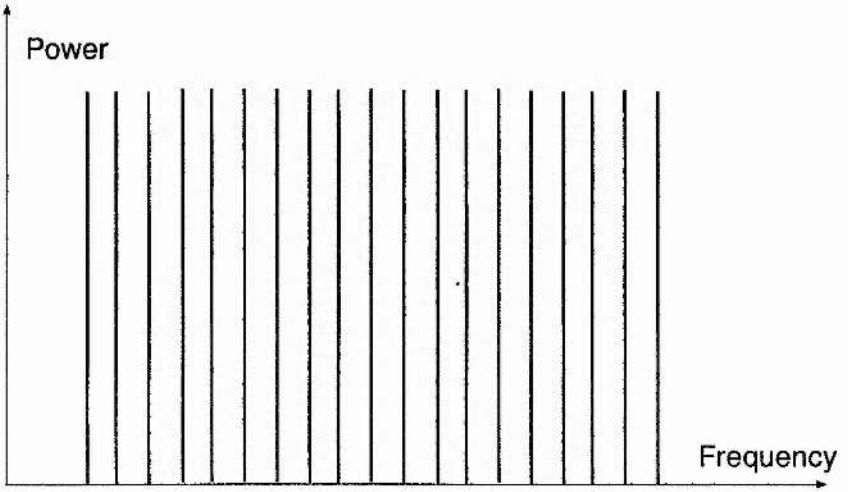


Fig. 6·5 : Ideal Frequency Hopped Spectrum

In practice of course, the spectrum will be composed of the desired frequencies plus sidebands and spuriæ generated as by-products of the imperfect (non-instantaneous) hopping process. A frequency hopping system is shown schematically in Fig. 6·6.

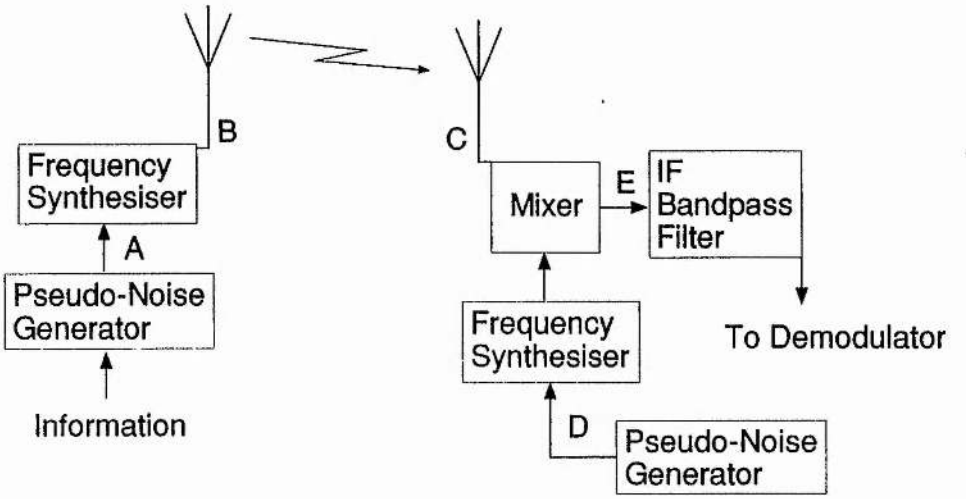


Fig. 6·6 : Frequency Hopping System

In the receiver, the signal is mixed with a locally generated reference from another synthesiser (at a fixed frequency offset) driven by an identical PN generator. When the signal and reference are synchronised, the carrier is recovered and bandpass filtered. The information can then be demodulated from the carrier. The waveforms corresponding to the points A to E above are shown in the following figure.

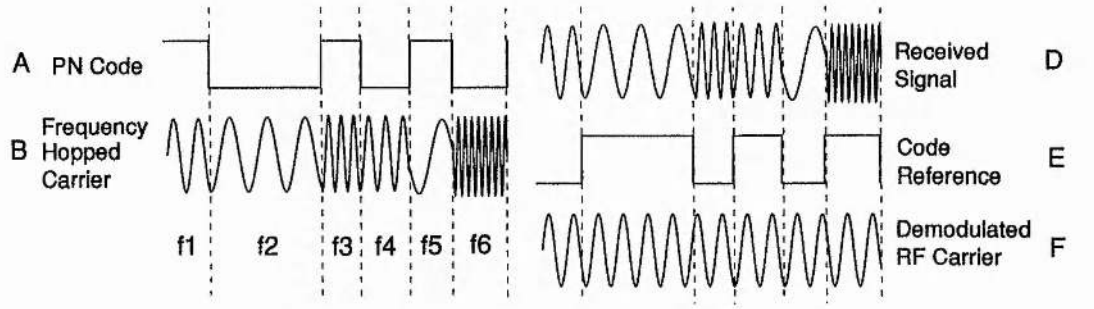


Fig. 6·7 : Frequency Hopping Waveforms

Any interfering signal is spread amongst the N channels by the action of the frequency hopped local reference and is rejected by the IF filter.

It is worth noting that whilst Fig. 6·5 shows well separated channels for clarity, channels can be contiguous or even overlapping to make more efficient use of bandwidth. For a system with contiguous channels the process gain is, as defined previously, the ratio of the rf to information bandwidths. However, for non-contiguous channels it is more accurate to use the expression

$$G_P = \text{number of available frequency choices} = N \qquad 6.6$$

The hop rate is determined by a number of factors:-

- the information rate
- the amount of redundancy employed
- the distance to the nearest potential interferer

Clearly, the bandwidth of the information signal sets the minimum hop rate, for which one hop per bit of information is the absolute minimum. Interference in a single channel can be overcome by sending more than one frequency per bit. This redundancy increases the hop rate but reduces the bit error rate (BER). For example[5], a 1000 channel system would have a BER of 1×10^{-3} with a single frequency jammer, whereas the same system with a redundancy of 3 hops per bit would have a BER of about 1×10^{-6}. This is a significant improvement in BER for a small increase in hop rate.

An additional problem is that of interference of the same frequency as the signal but of different phase. This can be caused by multipath[6] or by a deliberate jammer who retransmits the signal he has received from you within one hop duration. This effect then becomes a function of the distance and the offset angle between the receiver and the source of interference. This can be overcome by hopping at a rate greater than the speed at which the jammer can respond. For fixed communications links this may be quantifiable but for mobile links there is no alternative but to hop as quickly as possible.

6.2.3 Other Forms

There are a number of other schemes for implementing spectral spreading such as time hopping, pulsed FM (chirp), and hybrid methods like frequency hopping direct sequence, time-frequency hopping or time hopping direct sequence. They all have their peculiarities, advantages and drawbacks but space does not permit further discussion. For further details see Dixon (as in Ref. 3).

This chapter is intended to serve as an introduction to the main principles underlying spread spectrum techniques and as such cannot enter into much detail on the methods with which particular systems are engineered. Balanced modulators and frequency synthesisers are taken as ideal components for the sake of simplicity but I will now mention how the information is incorporated into the spread signal.

Analog information can be used to modulate the carrier before spreading but this suffers from a number of disadvantages. In a direct sequence system, the whole sinc-squared spectrum is deviated as the information frequency modulates the carrier. If this signal is then squared a component at twice the carrier frequency is produced which bears all the information in a classical FM manner. Therefore, such a system cannot offer privacy against all but the most casual of listeners.

When an FM carrier is used in a frequency hopped system, the two processes interact with each other and a hop from one frequency to another is demodulated as pulsed noise which is clearly undesirable.

Hence it is more usual to digitize the data first and then combine it with the code sequence. This can either be done by modulating the code clock or by modifying the code sequence in some way. Clock rate modulation is not usually used because any phase slippage between the received and local clocks would degrade performance.

The most useful method for DS systems is code modification. The information is embedded in the code so it can only be deciphered

by someone who has knowledge of that code. Typically, the data are re-clocked at the code clock rate and then mod-2 added to (XOR-ed with) the code sequence.

Frequency hopping systems use similar methods using logical operations on the code sequence to choose channels. It is worth noting that a FH system is inherently capable of protecting its internal code sequence to a higher degree than can a DS system. This is because the DS system modulates the carrier with the code and thus is directly observable, whilst in a FH system the code is not transmitted itself. Instead a frequency representing a code vector is sent.

6.3 Coding for Spread Spectrum applications

The capability of a spread spectrum system is determined by the properties of the pseudo-random code used to spread the signal in frequency space. The whole operability of these systems relies on using codes which are noise-like in their behaviour and yet deterministic. They provide protection against interference through the mechanism of process gain and they prevent eavesdropping by listeners who have no knowledge of the code in use.

There are many types of code in use with particular features which are suitable for particular applications. The most fundamental type of codes are known as maximal length codes.

6.3.1 Maximal Length Codes

Maximal length codes are the longest codes which can be generated by a digital shift register (or delay element) of a given length. A shift register generator (SRG) consists of a shift register and appropriate logic which feeds a logical combination of the state of two or more stages back to its input. For an n-bit register there are $2^n - 1$ states in a maximal length sequence[7].

Whilst maximal length sequences can be generated using

multiple tap feedback networks, a simple SRG with one mod-2 adder (XOR-gate) in the feedback network is the fastest possible configuration. Nonlinear networks do not produce maximal length sequences so they will not be discussed here.

Maximal length sequences are suitable for spread spectrum applications for a number of reasons:

- The number of 1's and 0's in the sequence is balanced allowing carrier suppression as discussed in section 6.2.1.
- Every possible state, bar one, of an n-bit register occurs once and only once in a complete code cycle. This provides $2^n - 1$ possible states which can be used to control uniquely a set of channels in a frequency synthesiser for frequency hopping. For the case of the XOR-gate feedback, the exception is the state of all zeros which is not permitted since it is self perpetuating.
- The autocorrelation is -1 for all offsets except for -1, 0 and +1 chip offsets, where the correlation varies linearly from -1 to $2^n - 1$ and back to -1 in a triangular peak - Fig. 6·8. This is calculated from a bitwise comparison of the code sequence with an offset copy. The number of agreements minus the number of disagreements equals the autocorrelation. This property is exploited in the receiver to determine when the incoming signal and local reference are in synchronism. Receiver correlation and synchronisation are dealt with later.

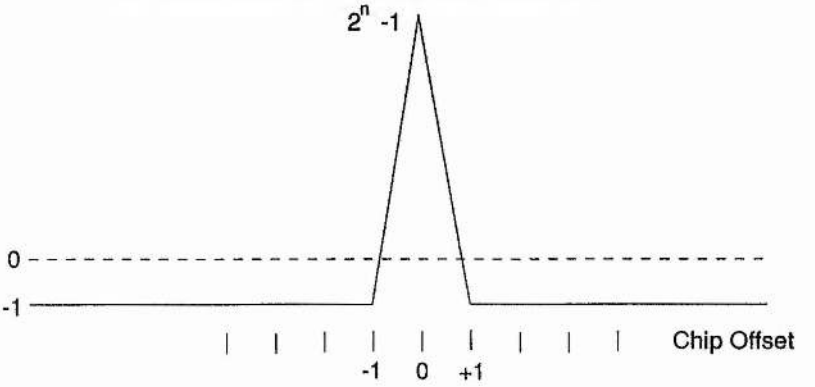


Fig. 6·8 : Autocorrelation of a PN sequence

Whilst autocorrelation is a measure of the self-similarity of a signal, crosscorrelation measures the similarity between one code sequence and another. This feature is exploited in multiple-access applications where a large number of users need different codes which will not interfere with each other.

Auto- and crosscorrelation are assumed to cover the complete code sequence. If they do not, subsets of the sequence may give rise to additional correlation peaks due to localised similarities. Furthermore, non-maximal length codes will also generate subsidiary correlation peaks and are thus unsuitable for spectral spreading.

One important point is that maximal length sequences can be deciphered. If the length of the register is known, then it is sufficient to collect $2n + 1$ sequential chips and be able to generate an identical sequence in its entirety[8]. Clearly long codes are better than short ones in this respect and the listener needs to know (or accurately estimate) the shift register length, but maximal length codes do not provide real security. Nonlinear codes must be employed for true message encryption.

6.3.2 Composite Codes[9]

Maximal length codes can be combined to produce other maximal length codes with interesting properties.

The Gold codes are generated by the mod-2 addition of a pair of n-bit maximal length sequences. The combination produces $2^n - 1$ new maximal length codes in addition to the two original ones. The cross-correlation properties between the codes is bounded and uniform so they are particularly suitable for multiple-access applications.

The JPL codes are formed from the mod-2 addition of two or more maximal length sequences whose length are relatively prime to one another. The result is a very long maximal length code which is suitable for accurate ranging over very long distances. The repeat

time of a code determines the maximum range at which the received signal can be unambiguously identified[10]. They are relatively simple to generate since only a few short shift register stages are needed. Additionally, synchronisation can be accomplished by separate operations on the (short) component codes and this permits fast synchronisation of the whole code.

6.4 Correlation and Demodulation

A spread spectrum receiver must perform two functions — de-spreading the incoming signal and demodulating the information. The latter function is undertaken by conventional techniques which are well documented. Direct sequence receivers tend to use coherent detectors such as the phase-lock loop or the Costas loop which work very well once the carrier frequency has been defined. However, these methods are not suitable in frequency hopping receivers since the loop would have to find lock every time the synthesiser changed frequency. Therefore, most FH systems use a much simpler scheme using an envelope detector and integrate-and-dump circuit. The rest of this section concentrates on methods of signal despreading and correlation.

The simplest way of correlating a PSK signal is to use another double balanced mixer like that used in the transmitter - known as an in-line correlator. This approach is by far the simplest but it is very poor at rejecting narrowband interference at the carrier frequency. Such signals tend to pass straight through and the receiver becomes very susceptible to jammers of this type.

In a heterodyne correlator, a DBM is again used but the local reference is at a constant frequency offset which produces a constant IF. In this way, direct feedthrough of the carrier frequency is not possible. Both DS and FH systems use heterodyne correlators for this reason.

The output of the correlator is the desired signal plus noise. The noise comes from the crosscorrelation of the local reference with

any unwanted signals. These unwanted signals arise from a number of sources. Thermal noise in the electronics and general rf noise from the transmission link contribute to the background noise level at the correlator output. Deliberate jamming signals are spread by the receiver and can constitute a major source of received noise. Additionally, if the timing of the received and local codes is offset, either before synchronisation is achieved or when synchronisation is imperfect, part of the desired signal is converted to pseudo-noise.

In DS receivers, lack of carrier balance and code modulation imbalance have been mentioned as being undesirable. Spurious components like these look like cw jamming signals and can seriously degrade system performance. In general, the carrier suppression of a balanced mixer should be at least equal to the process gain in the overall system[11]. It is possible to use tandem balanced mixers to increase the carrier rejection. An input filter before the correlator is used to reduce the level of stationary signals at the signal nulls.

6.5 Synchronisation and Tracking

Until this point, perfect synchronisation between the signal and the local code reference in the receiver has been assumed. It is necessary in all spread spectrum systems because the code is the vital key. Synchronisation is in fact one of the hardest problems to tackle and there are many ways in which it is approached.

Firstly, the code phase must be resolved to better than one chip in order that the despreading process can work correctly and secondly, the carrier frequency must be resolved accurately enough so that the despread signal fits within the post-correlation filter and demodulator.

Frequency stability of the reference alone is not sufficient as it cannot combat Doppler shifts in the received signal. Therefore, some active method of acquiring synchronisation and subsequently maintaining it is essential.

The basic method for acquiring synch is to use a sliding correlator in which the two codes are slipped in phase until coincidence is reached. However, one needs to cover the entire length of the code to cover all possible offsets and this must be done in a reasonable time. The rate of searching is limited by the post-correlation bandwidth. Clearly, this method is of limited use with long codes so there are a number of enhancements which can be made.

A preamble can be attached to the start of the code which is a short-cycle sequence designed to facilitate synchronisation. However, this approach is vulnerable to false correlation and reproduction by a jammer. Similarly, a reference signal can be transmitted along with the signal but this has no protection against interception.

Since the hop rate of a FH system is relatively low, some DS systems use a FH sliding correlator synchronisation subsystem before information transmission commences.

It is also possible to design a local code generator which extracts the incoming code and then "races ahead" to catch up with it. This is known as sequential estimation and can result in very fast acquisition.

Some of the fastest techniques use surface acoustic wave (SAW) devices as matched filters or convolvers to identify the signal-to-reference code phase offset. Parallel matched filters or serial-parallel hybrid architectures achieve a good compromise between speed of operation and bandwidth[12].

To correctly identify a correlation, a decision has to be made whether or not it is above a certain threshold level. The peak appears in a background of noise and the threshold must be set depending on what is an acceptable false alarm rate and probability of detection for the particular application.

Finally, the receiver then has to *stay* in synchronism with the signal. This is usually performed by a tracking loop[13], which follows the drift of the correlation peak and adjusts the code clock rate to maintain zero phase offset.

6.6 Applications

The advantages of spread spectrum systems can be summarised as follows[1,3,10,14]:

- interference suppression at low input signal-to-noise ratios
- low probability of intercept
- message screening
- multiple access and selective calling
- high resolution ranging

These properties are being exploited in an increasing number of applications. This final section looks briefly at a few of them.

One of the most rapidly expanding fields in which spread spectrum techniques are finding use is that of personal communication systems (PCS). Recent frequency allocations in the United States have encouraged the development of code division multiple access (CDMA) mobile telephone networks with message privacy[15,16] though at present this type of personal radio link is not licensed in the UK[17].

Wireless local area networks (WLANs) are a convenient way of providing connectivity for mobile or fixed computer systems. GEC have described a frequency hopped microwave radio system for this purpose which fits onto a single board just a few inches square[18].

The development of spread spectrum technology has largely been led by military interests and this will undoubtedly continue, though the resulting ideas and techniques often find subsequent commercial applications. The Global Positioning System (GPS) which is probably the best known example of a spread spectrum military system - see Dixon (as in Ref. 3) - has become a valuable tool for many civilians who require accurate navigational instrumentation.

REFERENCES

[1] Scholtz, R.A., "The Origins of Spread-Spectrum Communications", *IEEE Trans. Commun.*, **COM-30** (5), May 1982, pp. 822-854.
[2] Pickholtz, R.L., Schilling, D.L., 7 Milstein, L.B., "Theory of Spread-Spectrum Communications - A Tutorial", *IEEE Trans. Commun.*, **COM-30** (5), May 1982, pp. 855-884.
[3] Dixon, R.C., *Spread Spectrum Systems with Commercial Applications*, Wiley Interscience, 3rd Ed., 1994, pp. 10-15.
[4] ibid., Chapter 4.
[5] ibid., p. 39.
[6] Schumacher, P.M., "Model Propagation Effects in Spread-Spectrum Systems", *Microwaves & RF*, July 1993, pp. 98-102.
[7] Horowitz, P. & Hill, W., *The Art of Electronics*, Cambridge University Press, 1st Ed. 1988.
[8] Dixon, R.C., *Spread Spectrum Systems with Commercial Applications*, Wiley Interscience, 3rd Ed., 1994, Appendix 5.
[9] ibid., pp. 85-91.
[10] Baier, P.W., & Pandit, M., "Spread Spectrum Communication Systems", *Adv. Electron. & Elec. Phys.*, **53**, pp. 209-267.
[11] Dixon, R.C., *Spread Spectrum Systems with Commercial Applications*, Wiley Interscience, 3rd Ed., 1994, p. 171.
[12] Povey, G.J.R. & Grant, P.M., "Simplified matched filter receiver designs for spread spectrum communications applications", *Electron. & Commun. Eng. Jnl*, April 1993, pp. 59-64.
[13] Pickholtz, R.L., Schilling, D.L., & Milstein, L.B., "Theory of Spread-Spectrum Communications - A Tutorial", *IEEE Trans. Commun.*, vol. **COM-30** (5), May 1982, pp. 855-884.
[14] Dixon, R.C., *Spread Spectrum Techniques*, IEEE Press, 1976.
[15] Leuenberger, K.J., "Digital Radio Systems Examined - Parts 1 & 2", *Microwave Syst. News & Commun. Tech.*, Jan. 1986, pp. 81-92 & Feb. 1986, pp. 131-143.
[16] Schilling, D.L., Pickholtz, R.L., & Milstein, L.B., "Spread spectrum goes commercial", *IEEE Spectrum*, Aug. 1990, pp. 40-45.
[17] Vincent, J., "Voice link over spread spectrum radio", *Electronics World & Wireless World*, Sept. 1993, pp. 716-721. & Oct. 1993, pp. 826-833.
[18] Williams, D.A., "A Frequency Hopping Microwave Radio System for Local Area Network Communications", *IEEE MTT-S Digest*, 1993, pp. 685-690.

Chapter Seven
The Beacon Detection System - Part 1
RF Demodulation and Correlation

7.1 Introduction

I have carried out research into a quasi-optical, spread spectrum millimetric beacon detection system. This work has been part of a CASE award[1] and is described in Chapters 7, 8 and 9. Chapter 7 deals with the spread spectrum processing involved, Chapter 8 discusses the technique of spatial interferometry and its application in the system and Chapter 9 presents a theoretical analysis of system performance.

In October 1994 I presented the techniques described in Chapters 7 and 8 as an Invited Keynote Speech entitled "A Quasi-Optical Receiver for detecting Spread Spectrum Millimetre Wave Beacons" at the 19th International Conference on Infrared and Millimeter Waves, in Sendai, Japan.

7.2 The Beacon Detection System

7.2.1 Introduction

The Beacon Detection System (BDS) is designed to provide authorised listeners (or 'subscribers') with a quick method of positively identifying radio beacons. Additionally, the BDS must meet a given level of angular resolution in order to discriminate correctly between multiple beacons in the field of view.

The system as described consists of two main parts. Firstly, the beacon is a pseudo-randomly frequency modulated millimetre wave Gunn oscillator. Secondly, the receiver is a quasi-optical circuit which provides both angular discrimination and demodulation of the beacon signal. Associated processing electronics decode the FM signal and determine the identity of the beacon under observation.

Since the BDS uses wideband, coded FM signals, it can be considered, in general terms, to be a spread spectrum system. The rigorous definition of Chapter 6 does not apply because this is not a communication system, in the usual sense of a two-way link. Here the frequency spreading code *is* the information itself. Spread spectrum techniques and terminology are applied hereafter.

7.2.2 Beacon Requirements

The beacon is a continuously transmitting source which is intended to be received by only those listeners who have knowledge of the coding key. Therefore, the signal power should be low enough to make it difficult to detect with simple radiometric receivers, and the pseudo-random encryption must spread the signal over a wide bandwidth, in such a way that the signal can only be correctly recovered and identified by an authorised subscriber.

The beacons used are W-band (75-110GHz) Gunn oscillators, frequency modulated with a pseudo-random, digitally generated analog waveform. Gunn diode bias tuning characteristics allow frequency spreading over many hundreds of megahertz.

For the present work, it is assumed that all beacons operate at nominally the same frequency and are encrypted with the same coding key. However, each beacon will instantaneously be at a different point in the coding sequence from any other beacon. That way, each beacon is unique in its time offset position in the pseudo-random sequence. Multiple codes could be a future development.

The beacons are identified by correlating the received signal with a known copy of the coding sequence. Correlation peaks will exist for any positions in the sequence at which the signal and stored reference are aligned in time. Since all the beacons run at different time offsets, each correlation peak uniquely identifies one beacon.

In practice, for a secure, subscriber-only beacon system, it is assumed that the pseudo-random sequence will be chosen to have

very good auto- and cross-correlation properties, a high resistance to being decoded, and a very long repeat time. Generation of such a code sequence could be varied constantly, making use of the time of day as an updatable 'seed' parameter, for example.

7.2.3 Receiver Front End Requirements

The receiver optics are required to handle the incoming spread spectrum millimetre wave beacon signals and perform two functions on them. Firstly, as will be discussed in the next chapter, the quasi-optical circuit must provide a method of discriminating between multiple beacons. At W-band, the required angular discrimination would require an impractically large receiving antenna for a single port system. Consequently, for a real system with a smaller antenna, the beamwidth alone is insufficient to meet the requirement for angular resolution. Therefore, the real receiver needs to perform some quasi-optical signal processing in order to give enhanced angular discrimination. This process is fulfilled by using a two-port spatial interferometer (also known as a phase-comparison monopulse receiver).

Secondly, the W-band signal is demodulated by passing it through a quasi-optical frequency discriminator which converts the frequency variations of the beacon signal into voltage variations. Assuming the beacon is a linear frequency modulator and the discriminator also operates linearly, the voltage output of the discriminator should be a version of the modulating signal applied to the beacon. This voltage is then sampled and correlated with the stored copy of the coding key. In practice, the demodulated signal will also contain noise and distortion components due to nonlinearities of both the modulator and demodulator.

7.2.4 Receiver Back End Requirements

The post-detection signal processing would ideally be done

with digital signal processing (DSP) hardware. The sampled data are processed by the DSP board which performs all the necessary computations of apodisation, correlation and threshold detection. For the work described here, however, no DSP board was available so all testing of the receiver was done by executing the calculations on computer. This approach works adequately well, if a little slowly.

7.3 The Frequency Discriminator as an RF Demodulator

7.3.1 Theory of Operation

In order that the pseudo-random code which is used to modulate the beacon can be recovered, the spread spectrum transmission must be demodulated. A demodulator produces a voltage proportional to the instantaneous frequency of the incoming waveform and as such tries to recover the modulating signal. The modulating signal bandwidth is very much smaller than the transmission bandwidth so the demodulated signal is then easy to manipulate in analog or digital electronics. Such a discriminator can be implemented quasi-optically using a Martin-Puplett Interferometer with the path difference set appropriately - see Fig. 7.1 below.

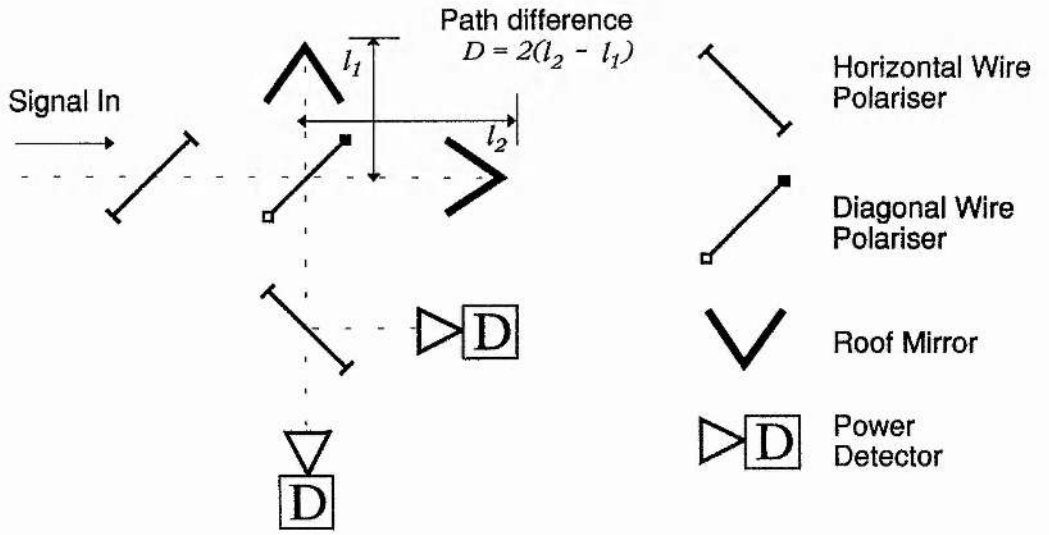


Fig. 7·1 : Frequency Discriminator

7.3.2 Circuit analysis of a Martin-Puplett Interferometer

Following the quasi-optical symbol conventions of Lesurf[2] it is possible to derive the response of a Martin-Puplett Interferometer with the use of Jones matrices. Each optical element is represented with a particular Jones matrix and even complex circuits can be analysed quite simply with a little matrix algebra.

The Jones matrices used here are defined as follows:-

Polariser	Transmission	Reflection
Wires vertical, $\theta = 0°$	$\mathbf{V_T}$ $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$	$\mathbf{V_R}$ $\begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$
Wires horizontal, $\theta = 90°$	$\mathbf{H_T}$ $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\mathbf{H_R}$ $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$
Wires at $\theta = +45°$	$\mathbf{P_T}$ $\frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$	$\mathbf{P_R}$ $\frac{1}{2}\begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}$
Wires at $\theta = -45°$	$\mathbf{N_T}$ $\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$\mathbf{N_R}$ $\frac{1}{2}\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}$
Distance travelled, d	$\mathbf{[d]}$ $\exp\left(-\frac{2\pi j d}{\lambda}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	
Roof mirror, roof line vertical	$\mathbf{R_V}$ $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	

Referring to the circuit shown overleaf, the electric fields at the output ports are derived by combining the Jones matrices of the optical elements acting on the input signal as it passes through the interferometer.

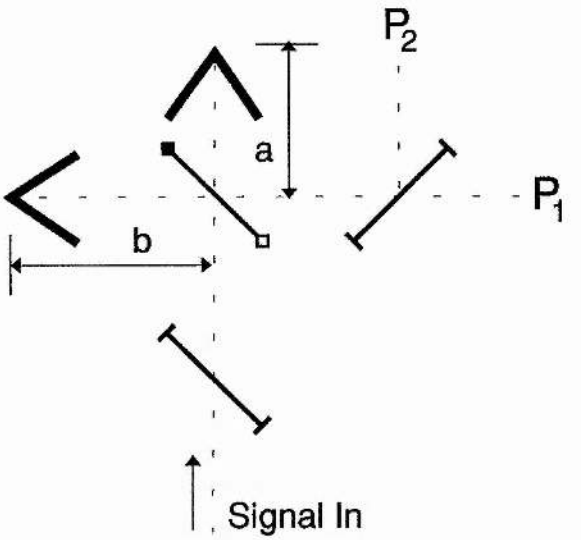


Fig. 7·2 : Martin-Puplett Interferometer Circuit

A vertically polarised input signal

$$E(t) = E_o \exp(-jA) \qquad \text{,where } A = \frac{2\pi f}{c} \tag{7.1}$$

can be expressed as a polarisation vector in the form

$$\boldsymbol{E} = (E_V,\ E_H) = \left(E_o \exp(-jA)\ ,0\right) \tag{7.2}$$

so that it can be manipulated with Jones matrices.

Hence, the field at P_1 is:-

$$\begin{aligned}\text{Field}(P_1) = {} & \mathbf{H_T N_R\,[a]\,R_V\,[a]\,P_T H_T}\left(E_o \exp(-jA)\ ,0\right) \\ & + \mathbf{H_T P_T\,[b]\,R_V\,[b]\,P_R H_T}\left(E_o \exp(-jA)\ ,0\right)\end{aligned} \tag{7.3}$$

and the field at P_2 is:-

$$\begin{aligned}\text{Field}(P_2) = {} & \mathbf{H_R N_R\,[a]\,R_V\,[a]\,P_T H_T}\left(E_o \exp(-jA)\ ,0\right) \\ & + \mathbf{H_R P_T\,[b]\,R_V\,[b]\,P_R H_T}\left(E_o \exp(-jA)\ ,0\right)\end{aligned} \tag{7.4}$$

which reduce to:-

$$\begin{aligned}\text{Field}(P_1) = {} & \tfrac{1}{2} \exp\left(\frac{-2\pi j 2a}{\lambda}\right)\left[-E_o \exp(-jA)\ ,0\right] \\ & + \tfrac{1}{2} \exp\left(\frac{-2\pi j 2b}{\lambda}\right)\left[-E_o \exp(-jA)\ ,0\right]\end{aligned} \tag{7.5}$$

and

$$\text{Field}(P_2) = \tfrac{1}{2}\exp\left(\frac{-2\pi j 2a}{\lambda}\right)\left[0, -E_o\exp(-jA)\right] + \tfrac{1}{2}\exp\left(\frac{-2\pi j 2b}{\lambda}\right)\left[0, E_o\exp(-jA)\right] \qquad 7.6$$

The power at the two ports can then be found as follows:-

$$P_1 = \left|\frac{E_o}{2}\right|^2\left|\exp\left\{-j\left(\frac{2\pi 2b}{\lambda}+A\right)\right\} + \exp\left\{-j\left(\frac{2\pi 2a}{\lambda}+A\right)\right\}\right|^2 \qquad 7.7$$

and

$$P_2 = \left|\frac{E_o}{2}\right|^2\left|\exp\left\{-j\left(\frac{2\pi 2b}{\lambda}+A\right)\right\} - \exp\left\{-j\left(\frac{2\pi 2a}{\lambda}+A\right)\right\}\right|^2 \qquad 7.8$$

Expand the exponentials in the form

$$\exp(-jZ) = \cos Z - j\sin Z \qquad 7.9$$

$$\exp(+jZ) = \cos Z + j\sin Z \qquad 7.10$$

and use the trigonometric identities

$$2\cos P\cos Q = \cos(P+Q) + \cos(P-Q) \qquad 7.11$$

$$2\sin P\sin Q = \cos(P-Q) - \cos(P+Q) \qquad 7.12$$

then substitute

$$P_o = E_o^2 \text{ and } D = 2(b-a) \qquad 7.13, 7.14$$

to yield the expressions for the power at the output ports:-

$$P_1 = \tfrac{1}{2}P_o\left[1 + \cos\left(\frac{2\pi D}{\lambda}\right)\right] \qquad 7.15$$

and

$$P_2 = \tfrac{1}{2}P_o\left[1 - \cos\left(\frac{2\pi D}{\lambda}\right)\right] \qquad 7.16$$

N.B. The power at each output port varies between zero and P_o.

7.3.3 The Martin-Puplett Interferometer as a Frequency Discriminator

For two power detectors of sensitivity, R, with incident power levels P_1 and P_2, the output voltages are

$$v_1 = \tfrac{1}{2}RP_o\left[1 + \cos\left(\frac{2\pi Df}{c}\right)\right] \qquad 7.17$$

and

$$v_2 = \tfrac{1}{2}RP_o\left[1 - \cos\left(\frac{2\pi Df}{c}\right)\right] \qquad 7.18$$

Passing these voltages through a differential amplifier of gain A gives an output of

$$v_{diff} = A(V_1 - V_2) = RAP_o \cos\left(\frac{2\pi Df}{c}\right) \qquad 7.19$$

Replacing the instantaneous frequency, f, with the instantaneous frequency deviation about nominal centre frequency i.e. $\Delta f = f - f_o$ one obtains

$$v_{diff} = RAP_o \sin\left(\frac{2\pi D\Delta f}{c}\right) \qquad 7.20$$

which contains the frequency variation term involving Δf.

Similarly, by passing the two voltages through a summing amplifier, also of gain A, we have

$$v_{sum} = A(V_2 + V_1) = RAP_o \qquad 7.21$$

Equation 7.21 shows that the sum signal contains only amplitude fluctuations and these can be removed from the frequency term by dividing eqn. 7.20 by eqn. 7.21 to give

$$v_d = \frac{v_{diff}}{v_{sum}} = \sin\left(\frac{2\pi D\Delta f}{c}\right) \qquad 7.22$$

Eqn 7.22 expresses the transfer characteristic of an ideal frequency discriminator. Frequency discrimination is still possible with only one detector (see eqns 7.17 & 7.18) but at the expense of halving the collected power and the inability to reject AM.

The path difference of the interferometer is set so that the discriminator output is nominally at one of the zero crossings of the

sinewave. Therefore, the discriminator tracks any variation in signal frequency by following the sinewave characteristic. This condition is met when $D = (n \pm \frac{1}{4})\lambda$, i.e. by setting the arm length difference to multiples of a quarter wavelength.

The figure below shows the characteristic curve of the frequency discriminator, for an arbitrary path difference, as the signal frequency is allowed to vary.

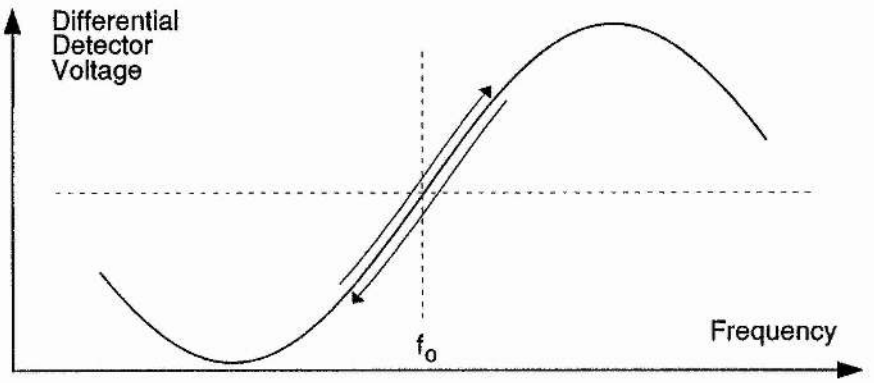


Fig. 7·3 : Discriminator Response with Signal Frequency

7.4 Linearity of Frequency Discriminator

Calculation of Error Bound on a Linear approximation to a Sinewave

A frequency discriminator which has a sinusoidal transfer characteristic can be considered as operating linearly only over a limited range. The degree to which the linear approximation holds is determined by what deviation from linear is considered acceptable for a given application. The deviation of a sinewave from a linear function can be computed in the following way.

The discriminator output is of the form

$$v_d = \sin\left(\frac{2\pi D\Delta f}{c}\right) \qquad 7.23$$

where D is the interferometer path difference, Δf is the instantaneous frequency deviation, and c is the velocity of light.

Now

$$\frac{dv_d}{dt} = \frac{2\pi D\Delta f}{c} \cdot \cos\left(\frac{2\pi D\Delta f}{c}\right) \qquad 7.24$$

so when the angle is zero (at a zero crossing) the last term becomes unity and the tangential linear approximation is just given by

$$v_l = \frac{2\pi D\Delta f}{c} \qquad 7.25$$

Define the error, e, of the sine function with respect to the linear function as

$$e \equiv \frac{v_l - v_d}{v_l} = \frac{\frac{2\pi D\Delta f}{c} - \sin\left(\frac{2\pi D\Delta f}{c}\right)}{\frac{2\pi D\Delta f}{c}}$$

$$\therefore \quad e = 1 - \text{sinc}\left(\frac{2\pi D\Delta f}{c}\right) \qquad 7.26$$

Plotting this function for a range of values of $D\Delta f$ shows the error bounds within which the FD should be operated. A graph of this function is shown overleaf.

For any given acceptable bound on the linearity, the product $D\Delta f$ sets the operating conditions of the discriminator. There is clearly a trade-off between bandwidth (Δf) and sensitivity.

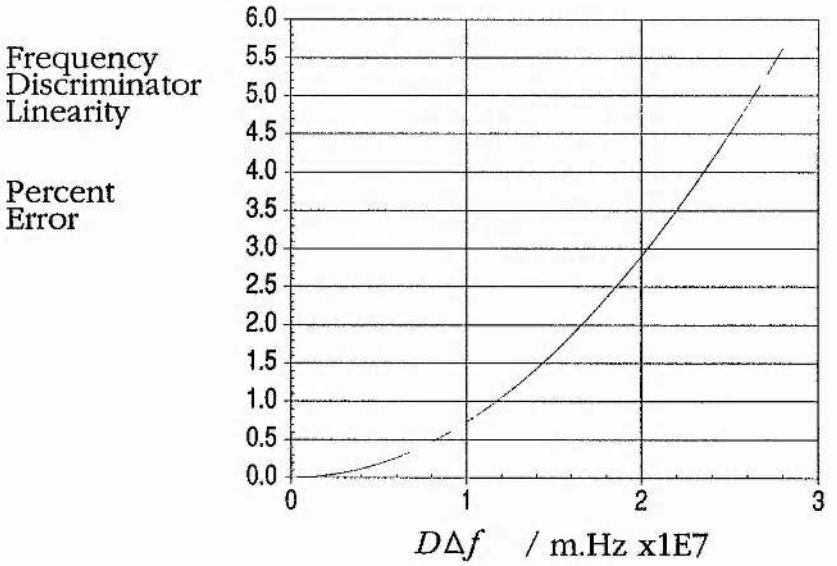


Fig. 7·4 : Linearity of frequency discriminator

These data are also presented below in terms of angular position on the discriminator sinewave.

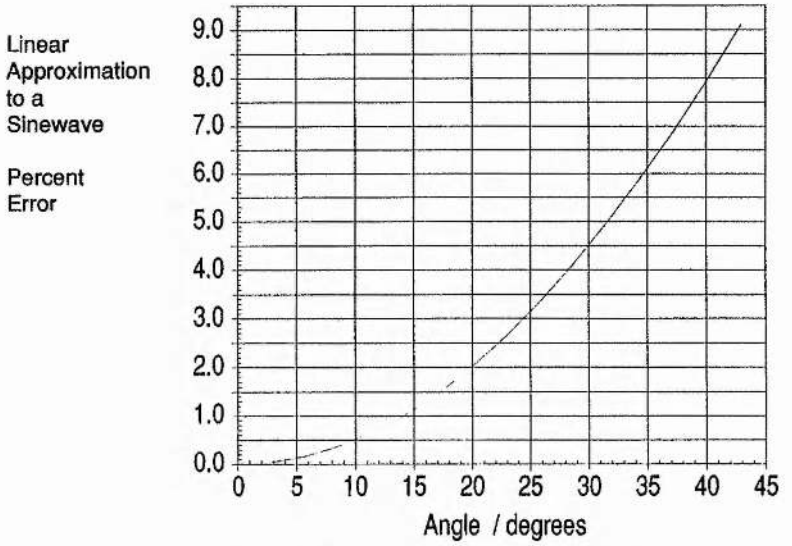


Fig. 7·5 : Linearity of a sinewave

If a 5% error is taken as an acceptable limit then $D\Delta f \leqslant 2.65E7$. Hence, if D = 40mm then $\Delta f \leqslant \pm 662.5$ MHz. From Fig. 7·5 one can see that the ±5% error bound on linearity of a sinewave occurs at ±32°, or at a value of ±0.53 on a unity amplitude sinewave. In other words, if the discriminator output for a modulated signal is kept within the half-amplitude range of the FD transfer characteristic then the nonlinearity error due to the FD itself will be less than 5%.

7.5 Optical Set Up

The interferometer is set up using an unmodulated oscillator by observing the outputs of the two detectors as the mirror position is varied - the detector signals should be sine-squared and cosine-squared patterns, varying from zero volts to a maximum corresponding to the output power of the oscillator - Fig. 7·6 - but in practice they are usually slightly distorted due to nonlinearity in the power detectors. The nominal 'zero crossings' of these signals are the positions of the mirror

which are suitable for frequency discrimination.

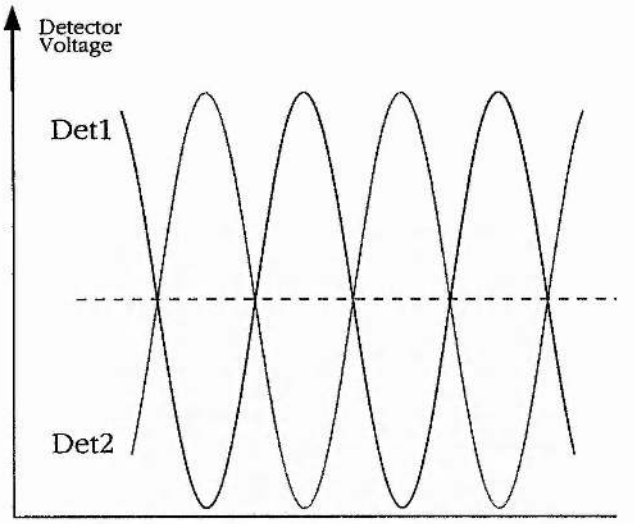


Fig. 7·6 : Detector Response with Path Difference

There is a tradeoff between the linearity bandwidth of the FD and the voltage sensitivity at the output - a greater D will give a steeper FD characteristic and hence greater voltage change for a given input frequency deviation but will only work linearly over a reduced bandwidth. The actual path difference chosen is therefore a compromise and I have usually set D to be the biggest value set by the 5% linearity condition. Note that a deviation of 5% is a peak value. The pseudo-random modulation has a roughly Gaussian distribution and so the majority of the waveform lies in the more linear region.

7.6 Correlation Software

In general, the cross-correlation of two functions $f(t)$ and $g(t)$ for a time delay of τ, is defined as

$$C_{fg}(\tau) = \int_{-\infty}^{+\infty} f(t).g(t+\tau)dt \qquad 7.27$$

$C_{fg}(\tau)$ can be used to compare two waveforms to measure their 'similarity'. Identical patterns will give a maximum correlation of unity when they are perfectly time-aligned, whilst dissimilar patterns will

correlate to relatively small values. In the context of the Beacon Detection System, the correlation function acts like a matched filter and is used to recover the coding key from the spread spectrum beacon transmission and identify the separate beacon signals.

For a pair of sampled datasets of length T, the function is modified to give

$$C_{fg}(\tau) = \frac{1}{T}\int_0^T f(t).g(t+\tau)\,dt \qquad 7.28$$

Hence the cross-correlations of the datasets $f(t)$ and $g(t)$ are computed for all possible time offsets within the data record length T.

Cross-correlation can be considered as a process in which two waveforms are multiplied together and integrated over time T. Therefore, the ability of a correlator to recover signals buried in noise is proportional to the integration time, which is normally the length of the data record - Fig. 7·7.

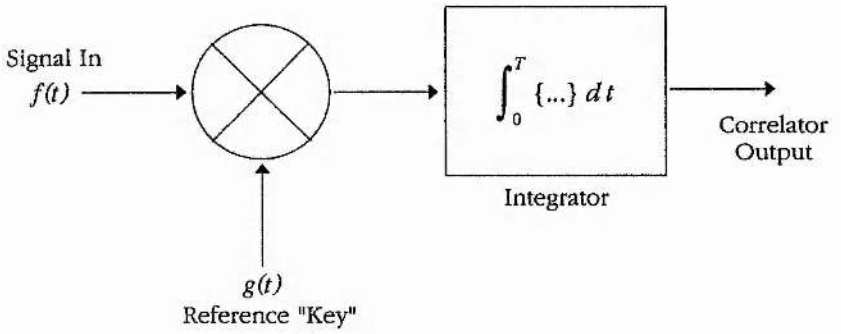


Fig. 7·7 : The correlation process

I have written software in BBC BASIC to compute correlations, based on the following techniques. The code is listed in Appendix 2.

Computationally, the correlation of two data sets is best executed using Fast Fourier Transforms (FFTs) in the following way:-

First take the FFT of each data set, then form the complex product of the two FFTs and finally take the inverse FFT of this product. The Real part of the inverse FFT will contain the correlation values for all possible time offsets in the data record and the Imaginary part should be all zero, or negligible at least.

Since any FFT assumes periodic data, the correlation datasets must be periodic over T, or at least start and finish on the same value.

If the data are repetitive within T, wraparound end-effects will cause the correlation to contain undesirable spurious peaks. To avoid this, the data set must be modified to be exactly one cycle long. The dataset can be adjusted using an interpolation routine based on the sinc function[3]. Each new data point is computed from a range of neighbouring original data points subject to weighting by the sinc function. Mathematically, the function is given as

$$x_n(t) = \sum_{i=n-K/2}^{n+K/2} x_i \,\mathrm{sinc}\left(\frac{\pi(t - t_i)}{\Delta t}\right) \qquad 7.29$$

where $x_n(t)$ is the n^{th} new sample value, x_i is the i^{th} original sample values, Δt is the original sample interval, t_i is the actual sample time of the i^{th} data point, and K is number of points used for the interpolation. In practice, K is set to 2 for speed of computation.

Adding the sinc interpolation to the correlation software removes spurious peaks and increases the signal-to-noise ratio on any given correlation - see Fig. 7·8.

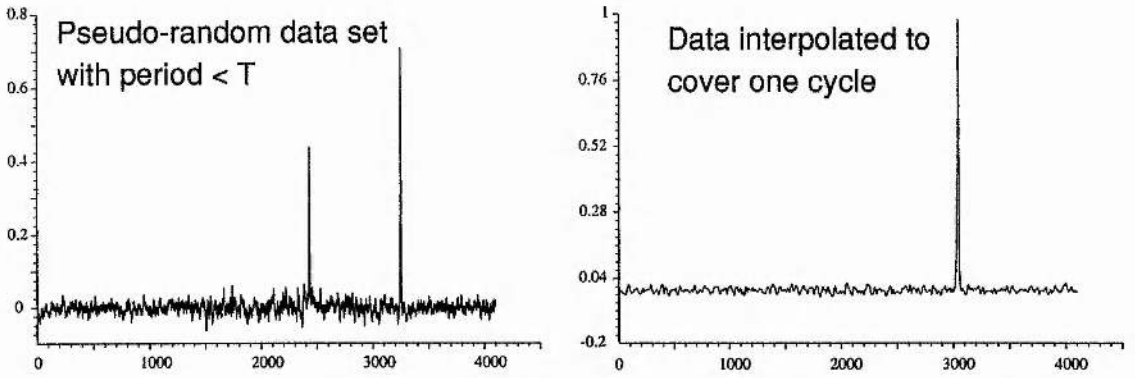


Fig. 7·8 : Removal of spurious correlations by interpolation of data

If the dataset starts and finishes on different values, resulting in a skewed baseline, the correlation baseline will be curved rather than flat, which makes threshold detection of the correlation peak very difficult. This can be overcome by apodising the data with a sinc function.

7.7 Experimental Results

The frequency discriminator and correlation detection software were tested extensively in the lab. The beacons are second harmonic GaAs Gunn oscillators operating at about 85GHz, frequency modulated by bias voltage modulation.

The pseudo-random modulation is generated digitally with a circuit based on an ALTERA EPM5016 device[4] programmed as a 15-bit maximal length shift register[5]. The sequence is maximal length so it has good correlation properties as described in Chapter 6. The pseudo-random code contains 32,767 bits and is clocked at 1MHz giving a sequence cycle time of 32.767ms (repeat frequency 30.5Hz). The digital noise is taken from the top bit of the shift register and converted to a bandlimited analog waveform of variable magnitude with a -3dB bandwidth of DC $\rightarrow$ 4.7kHz. It is the amplitude of the modulating waveform which determines the transmission bandwidth of the beacon, by exploiting the bias-to-frequency characteristics of the Gunn diode. Typically, the transmission bandwidth is 500 to 1000MHz.

The integration time of all the correlation measurements presented here was τ = 40.96 ms, which was the length of the sample dataset. The 4096 point datasets were acquired with a Thurlby DSA524 analog to digital convertor (ADC) sampling at 100kSa/s.

The discriminator is set to the correct path length using an unmodulated oscillator with the mirror positions being initially set according to the linearity condition. For a linearity limit of 5%, I set $D \approx 40$ mm, for a frequency deviation of $\Delta f = \pm 662.5$ MHz. The required position of the mirror is found by moving the mirror and monitoring the detector voltages as they trace out the curves of Fig. 7·6. Usually, the detector traces are distorted sinusoids and successive zero crossings alternate between two increments of mirror position. The average value of successive pairs do, however, correspond to within 100MHz of the measured signal frequency. The optimum mirror position is at the median of the two values which lie closest to 40mm.

The discriminator circuit is constructed from existing 70mm aluminium half-cubes on a half metre square baseplate with a 17.5mm pitch grid of dowel holes[6]. The square faces of the half-cubes hold lenses and the diagonal face holds a free-standing wire grid polariser. The lens apertures in the cube faces are 52mm in diameter.

Free-space Gaussian beams are coupled into receiving feedhorns[7] with high-density polyethylene (HDPE) planar-convex lenses[8]. The lenses employed are designed for use at 94GHz with a maximum throw of about 4 cubes (twice the distance from the lens to the output beamwaist) and are blazed for anti-reflection. The system was tested near 85GHz so the lenses would have been operating sub-optimally. The optical path through the discriminator is less than 4 cubes so the circuit does not suffer from significant truncation.

7.7.1 Correlation Results

On an optical baseplate, successful correlations were achieved for a single oscillator but when a second cw oscillator was added, the discriminator signal showed severe distortion. I examined the output for this case, but with triangle wave modulation. It appeared that the presence of the second free-running cw oscillator caused appreciable interference on the detected signal, typically like that shown in Fig. 7·9. Taking the difference of the two detector signals partially alleviates this common-mode problem.

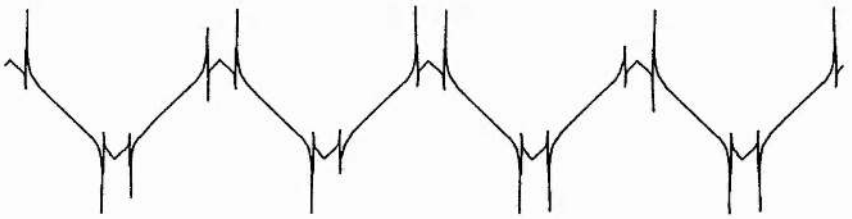


Fig. 7·9 : Injection Locking Spikes

I think the cause of this effect is due to an interaction between the two oscillators because they are running very close in

frequency. According to Dr Graham Smith[9], injection locking and increased FM noise sidebands are very typical in situations where two oscillators of similar frequency are coupled together quasi-optically. Very high levels of isolation would be needed between the oscillators to suppress these effects. In the BDS, this effect is unlikely to arise since individual beacons would be widely separated from the receiver and there would be negligible coupling of reflected signals back towards the sources.

I was able to measure directly the time offset between the discriminator signal and the reference signal using the cursors on the digital storage oscilloscope used to capture the waveforms. This was compared to the value obtained from the position of the correlation peak and it was found to be in agreement to within one interpolated sample interval (8μs) every time.

The ability of the discriminator to identify different beacons was tested using two beacons modulated with separate but identical noise signals. The transmission bandwidths were approximately equal. Once again, the time offset measurements made on the time record of the waveforms agreed with the correlations to within one sample interval. Fig. 7·10 shows a two-beacon correlation.

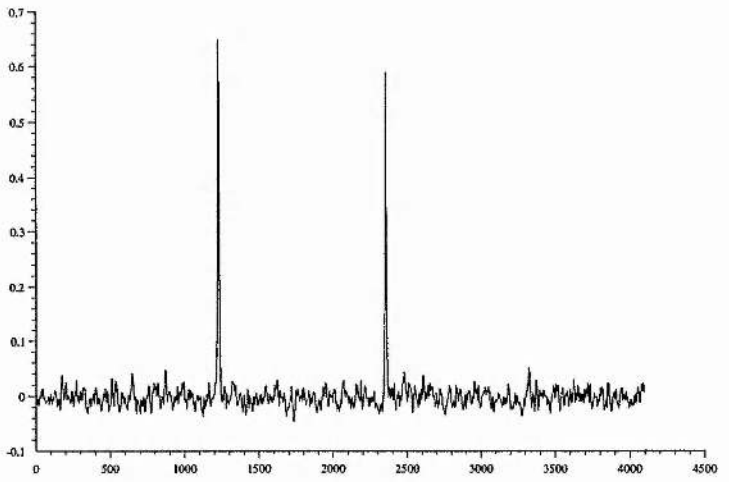


Fig. 7·10 : Correlation of two beacons, measured on optical bench

7.7.2 Minimum Detectable Signal Power

The finline power detectors used for these experiments have a quoted tangential sensitivity of -50dBm (10nW)[10]. Using slabs of attenuating material, I measured the correlation peaks for a range of received signal power. Adding a 10kHz low-pass filter to the discriminator output improved the correlations considerably by reducing the high frequency noise seen by the ADC. Most of this noise was due to the detector preamps - they are wide bandwidth and use OPA637 op-amps which are not particularly low noise. As a result, I was only able to achieve correlations down to power levels of the order of 100nW.

The detectors have a source impedance of about 50kΩ so a better choice of preamp would use a FET-input op-amp with a very low input current noise. I made new preamps using the AD743 BiFET op-amp which have a very good noise performance of $4nVHz^{-1/2}$ when referred to the input. With these improved preamps I recorded a correlation with an output signal-to-noise ratio, or SNR_o, (defined as the height of the peak with respect to the rms of the noise floor) of 17.6dB for a signal power of 23nW.

Most of the noise which is limiting these measurements comes from the detector because we are operating near to DC where the $\frac{1}{f}$-noise of the device is significant. By operating above the $\frac{1}{f}$-noise corner of the detector (about 1MHz for a silicon Schottky diode as used here), one could expect an improvement of 20dB.

A further improvement (of perhaps 5-10dB) may be possible by replacing the Si-Schottky diodes with planar doped barrier (PDB) devices[11]. They have a better tangential sensitivity (-57dBm) and a much lower $\frac{1}{f}$-noise corner than Si-Schottky diodes (about 20kHz).

Ultimately, the greatest sensitivity will be achieved by using heterodyne radiometers as power detecting elements. At W-band, Schottky diodes still offer better performance than PDBs when used as mixers. A typical mixer of comparable bandwidth to a power detector will be 50 to 60dB more sensitive. Of course, a radiometer

is considerably more complex than a power detector and the potential 30dB improvement outlined above, may be sufficient reason to retain power detectors for certain applications.

7.8 Testing the System at Range using Power Detectors

The receiver was tested at range by setting up a beacon on top of a platform in the field outside the lab window. The beacon was an oscillator with feedhorn positioned at ten metre intervals from the receiver which sat on the window sill. Only one detector was used for correlations in order to measure the power at the other output port with a Boonton 4220 power meter. The variation in SNR_o with range is shown in Fig. 7·11.

The SNR_o clearly drops off at range but using a wider transmission bandwidth, B_{RF}, improves matters due to the extra spread spectrum process gain.

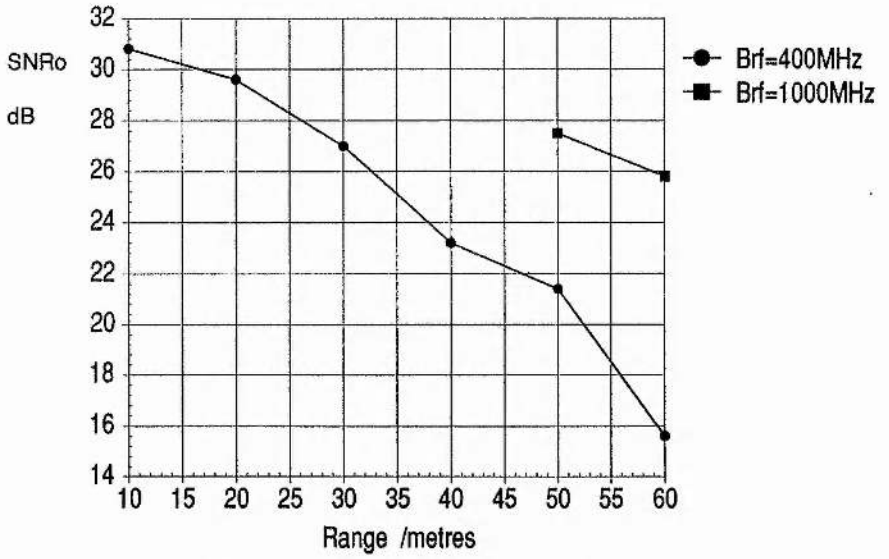


Fig. 7·11 : SNR_o vs Beacon Range using power detectors

To compare the measured received power with theory it is necessary to know the antenna gain of the receiver. This is determined from the antenna beam pattern.

With a cw source four metres across the lab from the receiver, the detector response was recorded with angular position - Fig. 7·12.

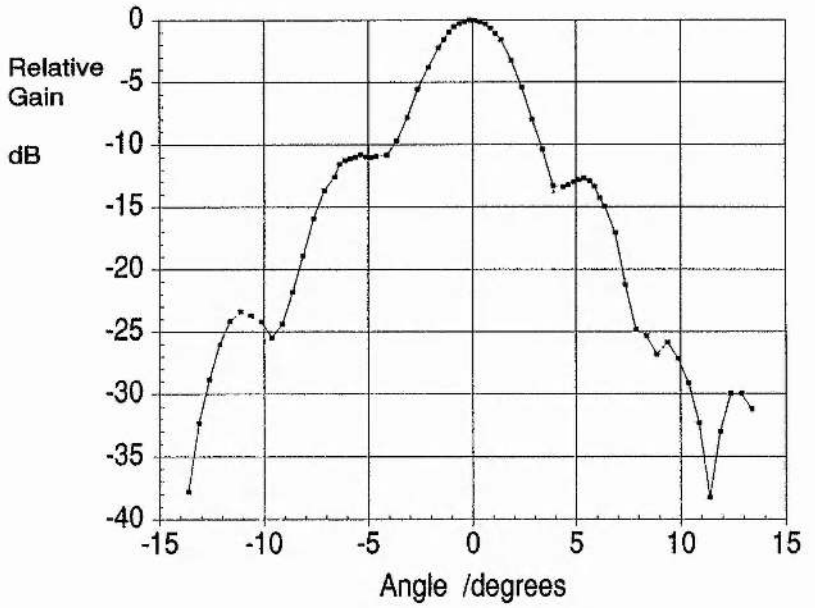


Fig. 7·12 : Frequency Discriminator Antenna Pattern

The antenna gain is a function of the $\frac{1}{e}$-power half-angle of the beam which in this case is $\theta_p = 2.2° = 0.0384$ rad.
The gain is given by

$$G = \frac{2}{{\theta_p}^2} \quad \text{,where } \theta_p \text{ is in radians} \qquad 7.30$$

As measured, $G_R = 1356 = 31.3$ dBi. The feedhorn is taken as having a gain of $G_T = 62 = 17.9$ dBi.

Using these values in the Link Gain equation

$$P_R = \frac{P_T G_T G_R}{\left(\frac{4\pi R}{\lambda}\right)^2} \qquad 7.31$$

to find the received power P_R, from a source of power $P_T = 20$ mW, and wavelength $\lambda = 3.49$ mm, at range R, the graph in Fig. 7·13 was constructed. The agreement of theory and experimental data is good.

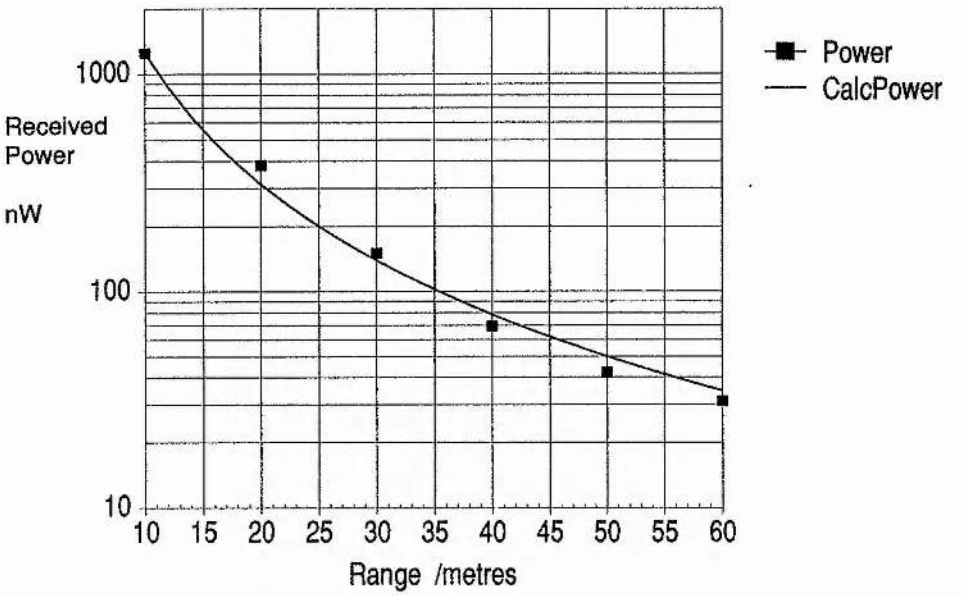


Fig. 7·13 : Link Gain theory and results for power detectors

7.9 Testing the System at Range using a Radiometer

To use the discriminator at longer range, the power detectors were insufficiently sensitive so I replaced them with a heterodyne radiometer system. The radiometer consisted of a single-ended mixer[12], battery bias supply, local oscillator fed via a 10-dB waveguide coupler, 70dB of IF amplification and an HP8472B crystal detector.

A series of measurements were taken at range for B_{RF} values of 400 and 1000MHz. Again, P_T = 20 mW but the feedhorn was removed from the oscillator so the radiation was coupled directly out from the open-ended waveguide. The best results are presented in Fig. 7·14.

Measurements at ranges greater than 200m were not conveniently possible in the vicinity of the laboratory.

It can be seen that the output signal-to-noise ratio does not degrade significantly at ranges up to 200 metres. Multipath effects and clutter from tree branches caused some variability in the results but the majority were all greater than SNR_o = 20dB.

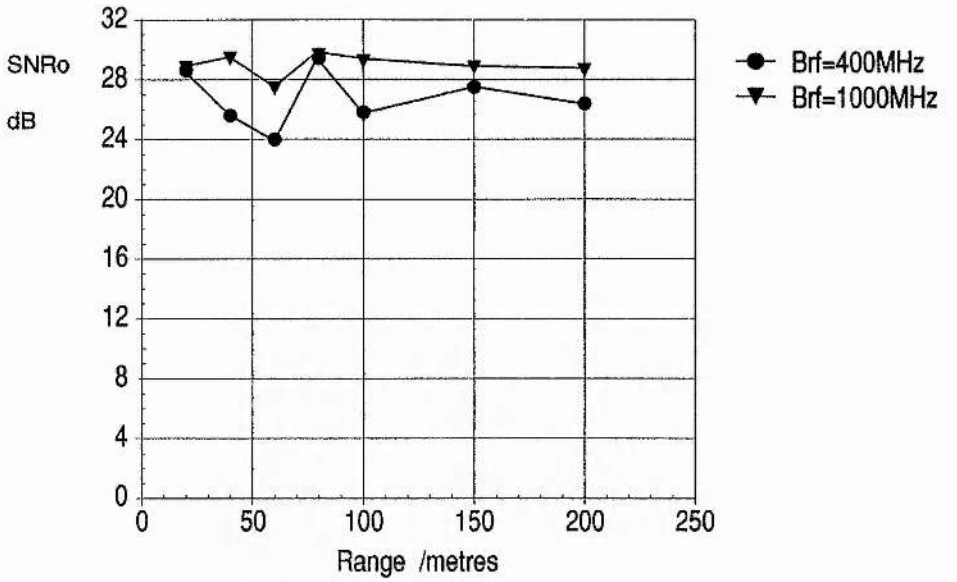


Fig. 7·14 : SNR_o vs Beacon Range using radiometer

In view of the massive increase in sensitivity which a mixer has over a power detector, correlations should be achievable at greatly increased ranges and should ultimately be limited only by the source power.

7.10 Conclusion

The techniques of quasi-optical frequency demodulation and digital correlation, and their application to the Beacon Detection System, have been introduced and described in theoretical terms.

Furthermore, an experimental quasi-optical rf demodulator with digital correlation, suitable for identifying spread spectrum millimetre wave beacons, has been successfully demonstrated. The spread spectrum advantage of process gain has been validated by demonstrating that wider bandwidth transmission improves performance. The sensitivity of the demodulator has been explored through using both power detectors and a heterodyne radiometer. When using power detectors, the experimental results are concurrent with theoretical predictions.

REFERENCES

[1] SERC CASE Award, ref no. 91700345, sponsored by DRA Malvern.
[2] Lesurf, J., *Millimetre-wave Optics, Devices & Systems*, Adam-Hilger/IOP, 1990.
[3] Lesurf, J.C.G., *Information & Measurement*, Lecture Notes, Physics Department, St. Andrews University, 1994.
[4] ALTERA Corp., 2610 Orchard Parkway, San Jose, CA 95134-2020
[5] Horowitz, P. & Hill, W., *The Art of Electronics*, Cambridge University Press, 1st Ed. 1988, pp. 437-442.
[6] Leeson, M.J., *The Application of Quasi-Optical Techniques to Millimetre Wave Radar*, PhD Thesis, St. Andrews, 1992.
[7] Thomas Keating Ltd., Station Mills, Billingshurst, West Sussex, RH14 9SH.
[8] Harvey, A.R., *A Quasi-Optical Millimetre Wave Complex Impedance Bridge*, Ph.D. Thesis, St.Andrews, 1990.
[9] Dr G.M. Smith, St. Andrews University, Private Communication.
[10] WDF-10 Finline Detector, Farran Technology Ltd., Ballincollig, Cork, Ireland.
[11] Kearney, M.J. & Dale, I., "GaAs planar doped barrier diodes for mixer and detector applications", *GEC Jnl. Res.*, **8** (1), 1990, pp. 1-12.
[12] WM-10 Single-Ended Mixer, Farran Technology Ltd., Ballincollig, Cork, Ireland.

Chapter Eight
The Beacon Detection System - Part 2
Spatial Interferometry and the Complete System

As was mentioned in the previous chapter, a degree of quasi-optical signal processing is necessary to achieve sub-beamwidth angular discrimination in the Beacon Detection System. The technique of spatial interferometry is employed for this reason in the BDS and will now be discussed in detail.

8.1 Principles of Two-Port Spatial Interferometry[1,2]

To provide sub-beamwidth angular resolution the receiver employs a two-port spatial interferometer which measures the difference between the relative phases of the signals incident on its two input antennas. The measured phase difference can be caused by either a change in range of the source or a change in its azimuth. Strictly speaking the 2-port system is unable to differentiate between the two effects but in practice it can be operated in conditions such that the phase difference due to changing range is negligible compared with that due to changing azimuth. This technique is also called *phase-comparison monopulse*[3]. Generally, it is assumed that the source is in the far field and that only the angle-dependence produces a significant phase difference. See Fig. 8·1 overleaf.

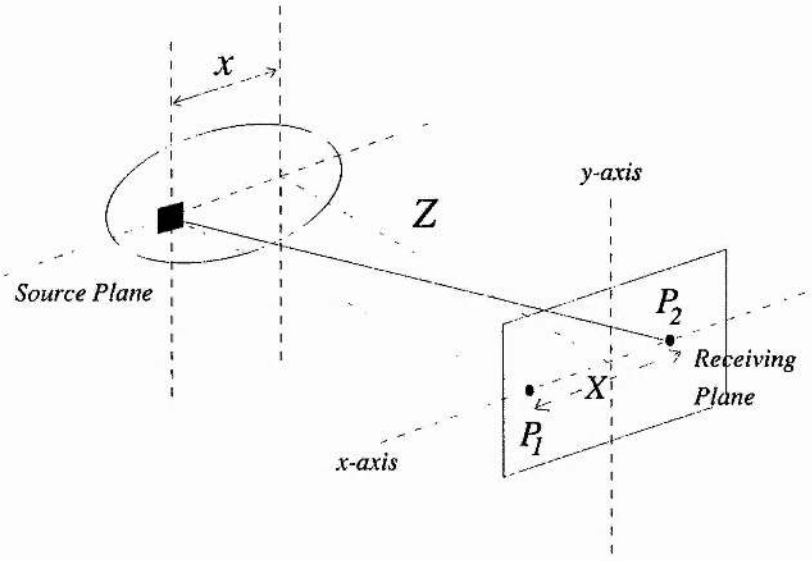


Fig. 8·1 : Two-port spatial interferometer geometry

8.1.1 Theory of Operation

For a system like that depicted above with two ports, P_1 and P_2, spaced at X, observing a source at range Z and azimuth x, the absolute phases of the waves incident on the two ports are

$$\phi_L = \frac{2\pi}{\lambda}\sqrt{Z^2 + \left(x - \frac{X}{2}\right)^2} + \phi\{t\} \qquad 8.1$$

$$\phi_R = \frac{2\pi}{\lambda}\sqrt{Z^2 + \left(x + \frac{X}{2}\right)^2} + \phi\{t\} \qquad 8.2$$

where ϕ[t] is a common phase term whose value depends on the instant of observation. Now both these phase terms contain an unknown number of nπ and it is impossible to solve for x, directly in terms of ϕ and Z. For large Z, the Binomial Approximation can be used to expand equations 8.1 and 8.2 to give

$$\phi_L = \frac{2\pi}{\lambda}\left[Z^2 + \left(x - \frac{X}{2}\right)^2\right]^{\frac{1}{2}} = \frac{2\pi Z}{\lambda}\left[1 + \frac{1}{2}\left(\frac{x - \frac{X}{2}}{Z}\right)^2 + \ldots\right] \qquad 8.3$$

and

$$\phi_R = \frac{2\pi Z}{\lambda}\left[1 + \frac{1}{2}\left(\frac{x + \frac{X}{2}}{Z}\right)^2 + \ldots\right] \qquad 8.4$$

which are combined to yield the following simple result:

$$\Phi = \phi_R - \phi_L = \frac{2\pi X}{\lambda} \cdot \frac{x}{Z} \qquad 8.5$$

Thus, the phase difference between the two ports is a linear function of the lateral offset at constant range and vice versa.

When either the source is on axis ($x = 0$) or the source is very far away ($Z \rightarrow \infty$) and at constant offset then the phase difference tends to zero. For any finite phase difference there is a locus of points (Z, x) which satisfy the general expression for $\Phi = \phi_R - \phi_L$ that traces out a hyperbola with foci at the positions of the two ports. The asymptotes of the hyperbola are given by the far-field approximation

$$x = \pm\frac{\lambda\Phi}{2\pi X}Z \quad ; \; \Phi \; constant \qquad 8.6$$

The angle of the asymptote is also the off-axis angle of the source.

Any particular value of Φ gives the gradient of the asymptote and this leads to the calculation of the eccentricity of the hyperbola and the coefficients with which the hyperbola can be generated. The difference between the hyperbola and the asymptote indicates an error which serves as a guide for when the asymptotic approximation is valid. In practice, the 1% error bound between hyperbola and asymptote is very close to the ports for a wide range of off-axis angles.

Strictly speaking, a source which moves along the path of the hyperbola will not cause a change in the measured phase difference. However, in reality this would require that the source knew the form of the hyperbola (implying knowledge of the observer's position, port spacing and operating wavelength) and would also require millimetre accuracy in following the hyperbola.

8.1.2 Circuit analysis of a Two-Port Spatial Interferometer

Following the same principles as used in Chapter 7, the two-

port spatial interferometer circuit can be analysed using Jones matrices. All the matrices necessary to analyse the circuit shown in Fig. 8·2 are defined in Chapter 7 with the exception of that for a plane mirror, which is:-

Plane mirror $\mathbf{M}$

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

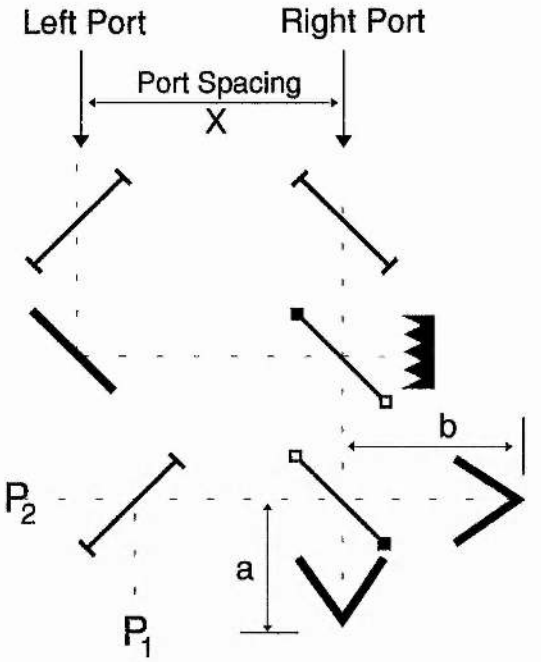


Fig. 8·2 : Two-Port Spatial Interferometer Circuit

The fields input at the left and right ports can be expressed as

$$\boldsymbol{E_L} = (E_o \exp(-jA), 0) \quad , \text{where } A = \tfrac{2\pi f}{c} + \phi_L \qquad 8.7$$

and

$$\boldsymbol{E_R} = (E_o \exp(-jB), 0) \quad , \text{where } B = \tfrac{2\pi f}{c} + \phi_R \qquad 8.8$$

where ϕ_L and ϕ_R are the absolute phases of the wavefront as it enters the two ports spaced at X.

Henceforth, the field at P_1 is:-

$$\begin{aligned} \text{Field}(P_1) = {} & \mathbf{H_R N_R [a] R_V [a] P_T P_R [X] M H_T} (E_o \exp(-jA), 0) \\ & + \mathbf{H_R P_T [b] R_V [b] P_R N_T H_T} (E_o \exp(-jB), 0) \end{aligned} \qquad 8.9$$

and the field at P_2 is:-

$$\text{Field}(P_2) = \mathbf{H_T N_R [a] R_V [a] P_T P_R [X] M H_T}\,(E_o \exp(-jA)\ ,0) + \mathbf{H_T P_T [b] R_V [b] P_R N_T H_T}\,(E_o \exp(-jB)\ ,0) \quad 8.10$$

which reduce to:-

$$\text{Field}(P_1) = \tfrac{1}{2}\exp\left(\frac{-2\pi j(2a+X)}{\lambda}\right)\left[0,\ -E_o\exp(-jA)\right] + \tfrac{1}{2}\exp\left(\frac{-2\pi j2b}{\lambda}\right)\left[0,\ E_o\exp(-jB)\right] \quad 8.11$$

and

$$\text{Field}(P_2) = \tfrac{1}{2}\exp\left(\frac{-2\pi j(2a+X)}{\lambda}\right)\left[-E_o\exp(-jA)\ ,0\right] + \tfrac{1}{2}\exp\left(\frac{-2\pi j2b}{\lambda}\right)\left[-E_o\exp(-jB)\ ,0\right] \quad 8.12$$

The power at the two output ports can be found as follows:-

$$P_1 = \left|\frac{E_o}{2}\right|^2\left|\exp\left\{-j\left(\frac{2\pi 2b}{\lambda}+B\right)\right\} - \exp\left\{-j\left(\frac{2\pi(2a+X)}{\lambda}+A\right)\right\}\right|^2 \quad 8.13$$

and

$$P_2 = \left|\frac{E_o}{2}\right|^2\left|\exp\left\{-j\left(\frac{2\pi 2b}{\lambda}+B\right)\right\} + \exp\left\{-j\left(\frac{2\pi(2a+X)}{\lambda}+A\right)\right\}\right|^2 \quad 8.14$$

Expand the exponentials in the form

$$\exp(-jZ) = \cos Z - j\sin Z \quad 8.15$$

$$\exp(+jZ) = \cos Z + j\sin Z \quad 8.16$$

and use the trigonometric identities

$$2\cos P\cos Q = \cos(P+Q) + \cos(P-Q) \quad 8.17$$

$$2\sin P\sin Q = \cos(P-Q) - \cos(P+Q) \quad 8.18$$

then substitute

$$P_o = E_o^2 \text{ and } D = 2(b - a) \qquad 8.19, 8.20$$

to yield the expressions for the power at the output ports:-

$$P_1 = \tfrac{1}{2}P_o\left[1 - \cos\left\{\frac{2\pi(D - X)}{\lambda} + \Phi\right\}\right] \qquad 8.21$$

and

$$P_2 = \tfrac{1}{2}P_o\left[1 + \cos\left\{\frac{2\pi(D - X)}{\lambda} + \Phi\right\}\right] \qquad 8.22$$

where $\Phi = \phi_R - \phi_L$ is the wavefront phase difference at the input ports.

N.B. The power at each output port varies between zero and P_o which is only half the total input power (which is $2P_o$, because we have two ports) since half the power is thrown away at the first diagonal polariser. If a 45° Faraday rotator was added after each input polariser, all the power would be collected at the output ports, thereby removing the factor of $\frac{1}{2}$ in equations 8.21 and 8.22.

For the interferometer described here, the path difference D, is set equal to X in order to cancel out the phase difference introduced by the finite port spacing, so the expressions simplify to

$$P_1 = \tfrac{1}{2}P_o\left[1 - \cos\left(\frac{2\pi X}{\lambda} \cdot \frac{x}{Z}\right)\right] \qquad 8.23$$

and

$$P_2 = \tfrac{1}{2}P_o\left[1 + \cos\left(\frac{2\pi X}{\lambda} \cdot \frac{x}{Z}\right)\right] \qquad 8.24$$

when Z is large and eqn. 8.5 is valid.

Hence, for fixed X and λ, the response of the spatial interferometer is determined by the source range Z and lateral offset x. However, on or very near axis ($x \approx 0$), the system is virtually

immune to changes in range, but is relatively sensitive to changes in azimuth. Therefore, the system as implemented provides sub-beamwidth determination of the azimuthal position of a source. The sensitivity is controlled by the port spacing X.

Now whilst the theoretical response is periodic (due to the cosine term), the real system is also subject to the finite beamwidth of the receiver antenna pattern. This has the effect of overlaying a Gaussian envelope on the cosinusoidal response of the interferometer i.e.

$$P_1, P_2 = \exp\left(-\left(\frac{\theta}{\theta_p}\right)^2\right) \cdot \frac{P_o}{2}\left[1 \mp \cos\left(\frac{2\pi X\theta}{\lambda}\right)\right] \qquad 8.25, 8.26$$

where $\theta = x/Z$ in the cosine term.

The resulting response versus θ is shown in Fig. 8·3 for the following values:- $X = 70$ mm, $\lambda = 3.49$ mm, $\theta_p = 2.25°$, $P_o = 1$

P_1 and P_2 are referred to as the Difference and Sum channels respectively, in keeping with monopulse radar terminology.

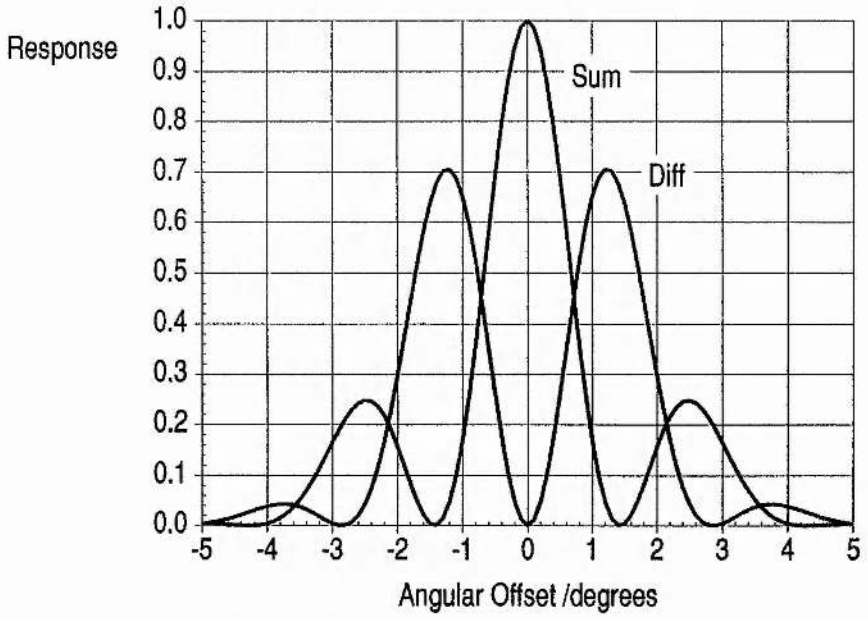


Fig. 8·3 : Theoretical Response of Spatial Interferometer

It is clearly seen in Fig. 8·3 that a signal source on or near axis will generate simultaneously a large response in the sum

channel and a null response in the difference channel. Only in this narrow angular region is the difference between the two channels so acute. Therefore, by making a relative measurement between the sum and difference channels we can determine the off-axis position of a source to a much greater accuracy than is possible with a single antenna beam.

8.1.3 Conventional Monopulse

Monopulse radar techniques are discussed in detail in the book by Leonov & Fomichev[4]. Conventional amplitude-comparison monopulse utilises a pair of coincident antennas which are squinted at some angle. The sum and difference antenna amplitude responses are used to determine the angle of the source signal. Spatial interferometry is also known as phase-comparison monopulse and is characterised by the use of a pair of parallel-pointing antennas which are deliberately spaced apart.

Both systems can be described as having a sum and a difference channel which are used for angle sensing - the sum channel is taken to mean that which has a maximum on axis, and the difference is taken as having a null on axis - and these conventions are used throughout this thesis.

8.2 Spatial Interferometer - Experimental Results

Angular Discrimination

The circuit as shown in Fig. 8·2 was set up with a port spacing of 70mm (one cube width) and the path difference also set carefully with a ruler to 70mm in such a way as to counteract the effect of the port spacing. The power at the two output ports was measured with a pair of power detectors. HDPE lenses were used as before, but in this case, the optical path length exceeds the maximum throw of 4 cubes so a certain amount of truncation is likely in the circuit. Time did not

allow for the design of a series of lenses that would optimise coupling.

A source with a feedhorn was placed 4 metres away at the same height as the receiver and pointed towards the input ports and the output response was measured with angular position. The receiver was mounted on a turntable inscribed with a vernier scale marked in increments of 5 minutes of arc.

Initially, the sum and difference patterns were asymmetric in height due to an overall skewing of the response. This was found to be due to unequal responses of the input ports. The pointing accuracy of the source affected the beam patterns considerably and great care had to be taken to align the source feedhorn with the receiver axis.

Once the antenna patterns of the left and right ports were equalised, the angular response measurements showed quite good agreement with theory.

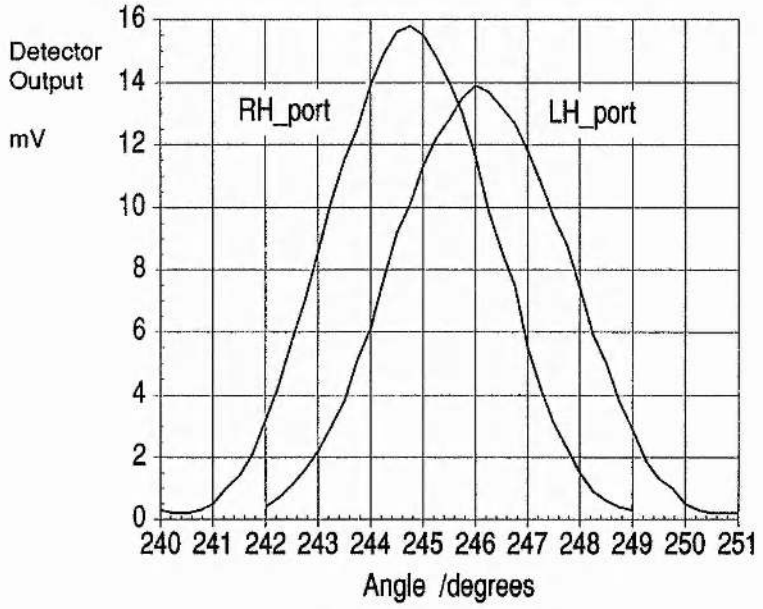


Fig. 8·4 : Antenna patterns of each port of spatial interferometer

The $\frac{1}{e}$-power beamwidth, θ_p, is taken as ±2.25°. From this, the output beamwaist can be calculated as being ω_o = 20 mm using

$$\omega_o = \frac{\lambda}{\sqrt{2}\pi\theta_p} \qquad 8.27$$

From this, one can calculate the antenna gain as being

G_R = 1295 = 31 dBi using

$$G_R = \frac{4\pi^2\omega_o^2}{\lambda^2} \qquad 8.28$$

The response of the spatial interferometer was measured and the results fitted with theoretically derived curves. The computed curves use the same parameter values as for Fig. 8·3, but also include a 0.78° skew angle between the two ports. Compared to Fig. 8·3, it can be seen that skewing the beams causes imperfect cancellation of the two signals and the minima do not reach zero. Also, since the detector sensitivities are unequal and uncalibrated, the D2 dataset has been scaled down by a factor of 1.2 to obtain the best fit with theory.

The measured data fit the computed curves very well,. The sum and difference data are quite symmetric and the peaks and troughs match the theory adequately. Whilst the region within ±3° is a very good fit, the extremes of the difference data set fall away more rapidly than predicted but this is not considered important. A skew angle of 0.78° is probably caused by misalignment in the interferometer arms.

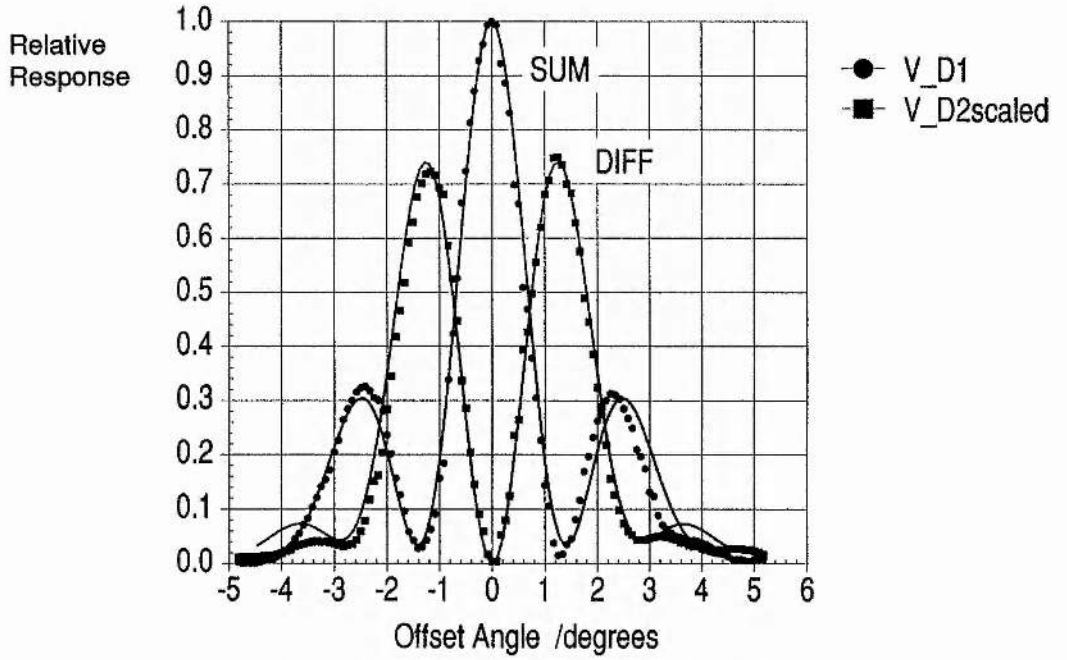


Fig. 8·5 : Spatial Interferometer Response - Linear Scale

It is instructive to look at the response on a log scale to observe

the depth of the null in the difference channel. The measured and scaled results are shown in a log plot in Fig. 8·6.

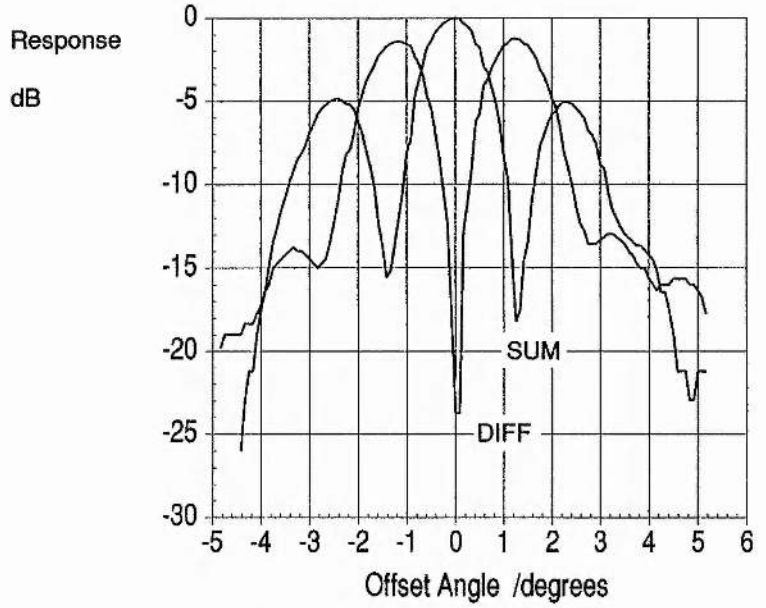


Fig. 8·6 : Spatial Interferometer Measured Response - Log Scale

One can clearly see that on axis, the difference channel rejects signals by 23dB with respect to the sum channel. The behaviour in the region of this null is shown in more detail in Fig. 8·7, which plots the difference between the curves plotted in Fig. 8·6.

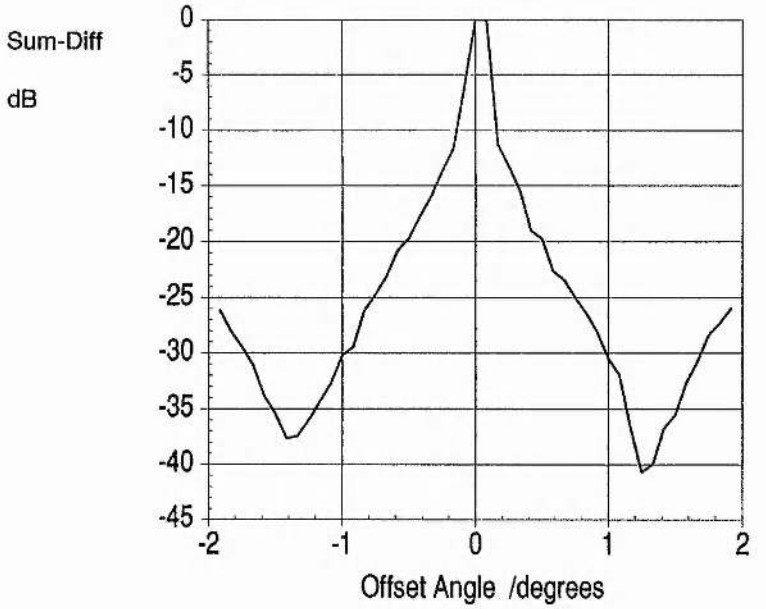


Fig. 8·7 : Detail of measured null rejection (Sum-Difference)

The deeper the null, the better the rejection of a signal in the difference channel with respect to the sum channel. At the -12dB level, the angular offset is approximately ±0.2° which is less than one tenth of the antenna beamwidth. This shows the ability of the spatial interferometer to discriminate very accurately between sources located close together. An angle of 0.2° corresponds to an azimuthal offset of 0.35 metres per 100m of range.

8.3 The Complete System

Theory of Operation

The complete system consists of a spatial interferometer with a frequency discriminator on both the sum and the difference channels. In this way, beacons on axis will give a correlation peak in the sum channel, but none in the difference channel. Thus if a correlation peak appears in the sum channel but is rejected in the difference channel then the beacon under observation must be on axis. For any beacons located elsewhere, they will produce a correlation peak in both the sum and difference channels. This is shown schematically in Fig. 8·8.

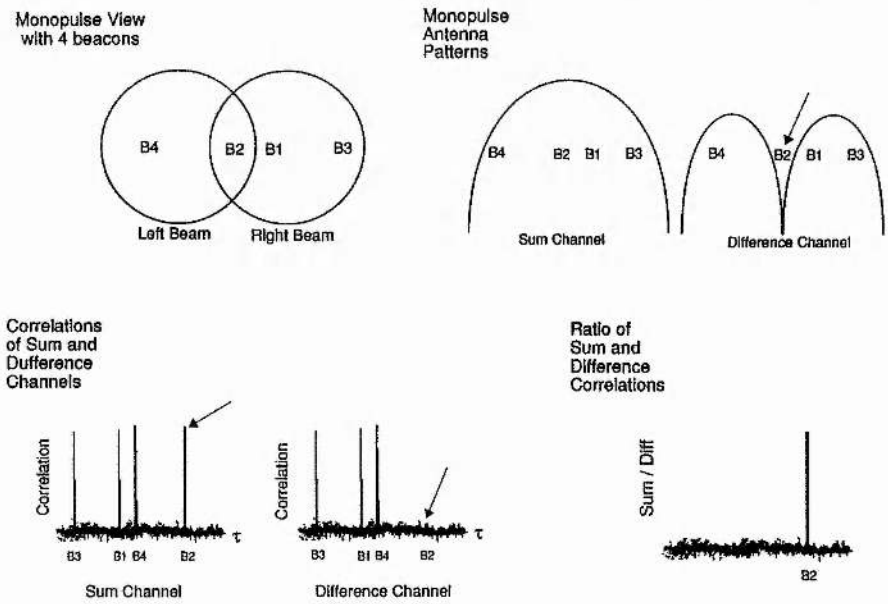


Fig. 8·8 : Operating principle of complete system

Only beacon B2 is on axis. The sum and difference patterns depicted are amplitude-comparison monopulse antenna patterns for simplicity. In this phase-comparison monopulse system, they will look like Fig. 8·6. All the beacons generate correlation peaks in the sum channel but only B2's difference channel correlation is suppressed. The ratio of the sum and difference channels displays the unique property of B2 - that it is positioned on axis. The quasi-optical circuit of the complete system is shown in the next figure and a photograph of the receiver optics is shown overleaf.

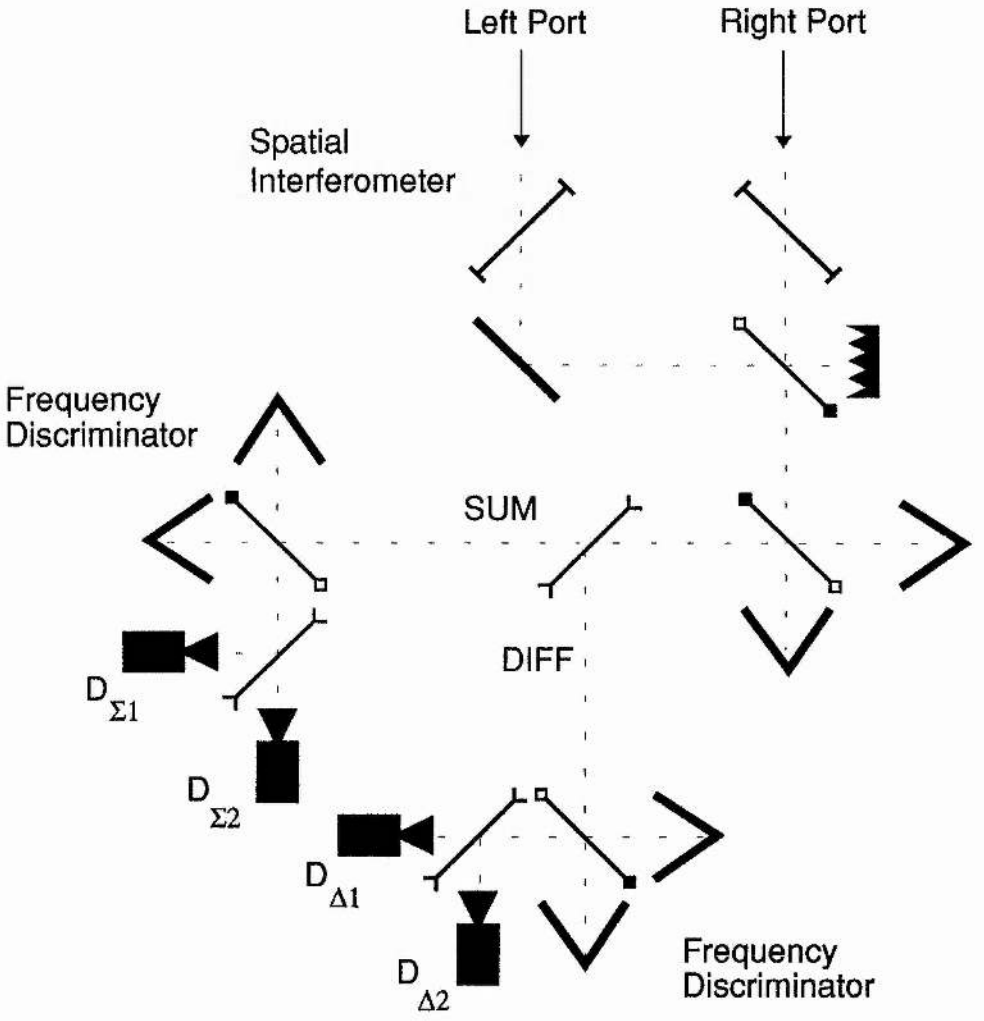
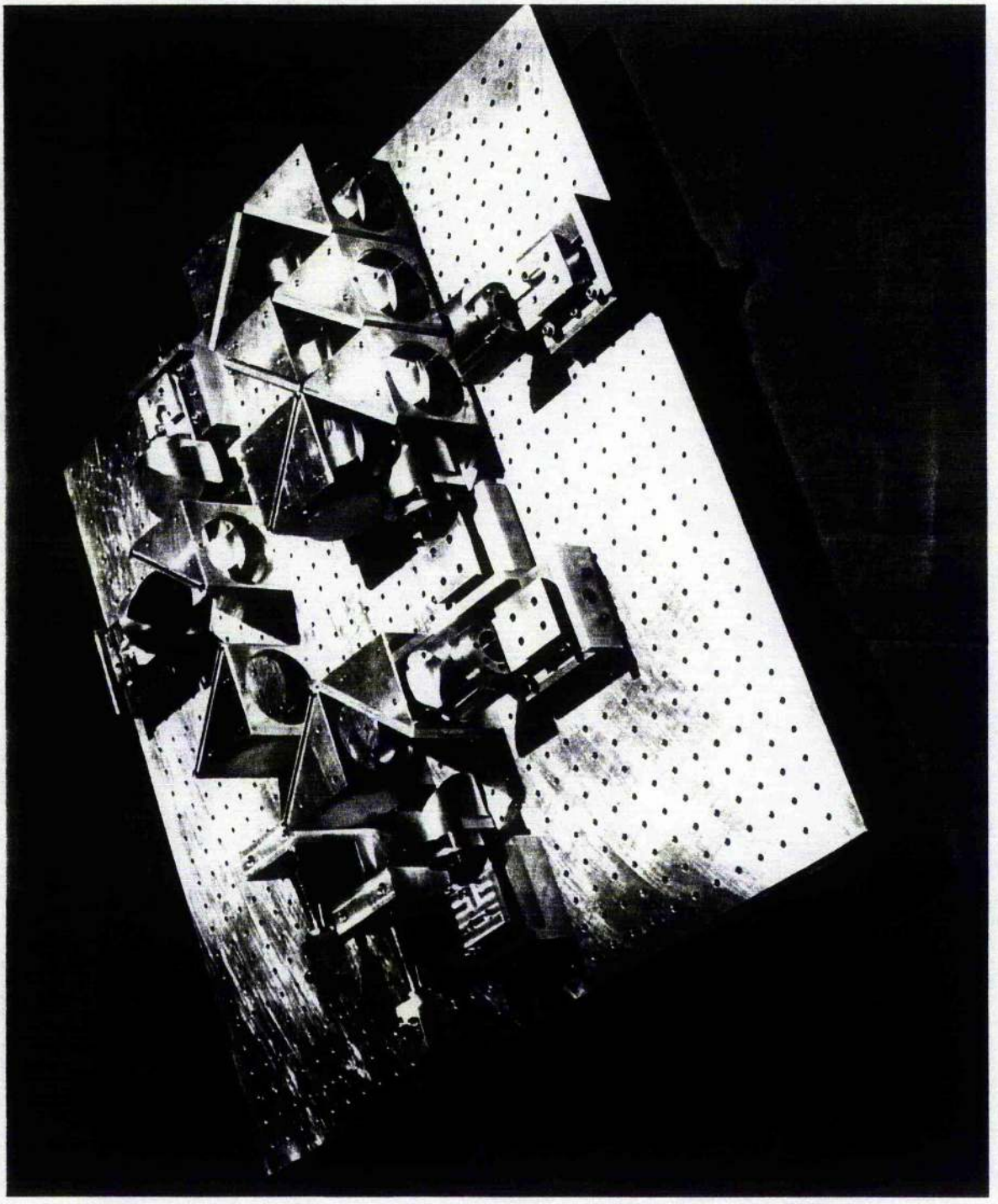


Fig. 8·9 : QO circuit of complete system



The Complete Quasi-Optical Receiver Circuit

8.4 The Complete System - Experimental Results

8.4.1 Angular Response

The response of the full system was measured in a similar way to that described for the spatial interferometer alone. The receiver was mounted on a turntable, 4 metres away from a source with a feedhorn, positioned on axis to the receiver. The spatial interferometer path length was set to 70mm as before and the frequency discriminators set to give maxima at the two detectors which were used, one per discriminator. Two long focal length lenses were added to the sum and difference outputs of the spatial interferometer to improve coupling and reduce truncation effects through the full circuit. The extra lenses almost doubled the collected power level but a further improvement would be achievable with a correctly designed set of lenses.

The antenna patterns of both input ports were measured at all four output ports and the results are displayed in Fig. 8·10.

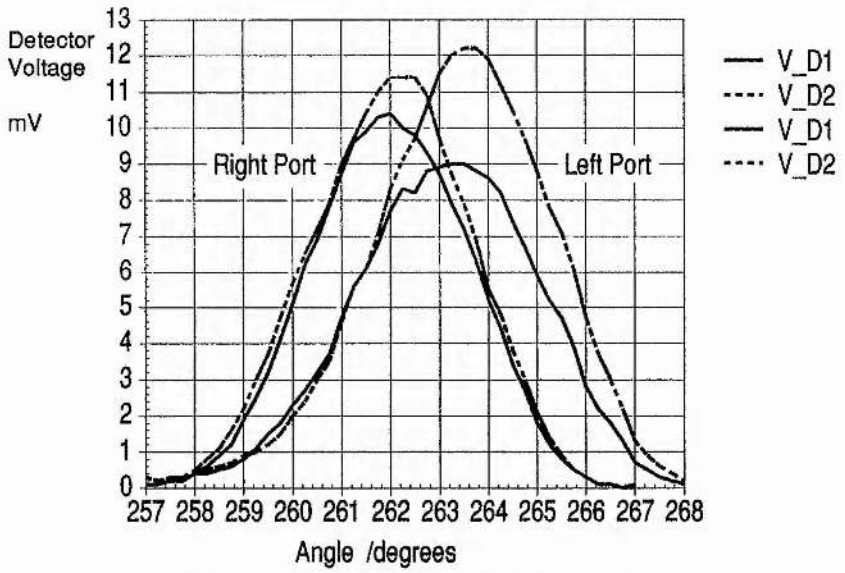


Fig. 8·10 : Antenna patterns of complete system

Whilst the patterns measured for the right hand port are consistent, those for the left hand port differ moderately. This is likely to be due to a combination of the unequal sensitivities of the

detectors (D2 has a greater sensitivity than D1) and different VSWR effects at the output ports. The mismatch was not considered as significant at this stage and time did not allow for further refinements to the optical set-up. The $\frac{1}{e}$-power beamwidth, θ_p, is taken as ±2.25°, giving an output beamwaist of ω_o = 20 mm and an antenna gain of G_R = 1295 = 31 dBi.

The frequency discriminators were then adjusted, as described in detail in the relevant section above, to sit at the half-way point of the discriminator sine curve.

The angular response of the system was measured using one power detector per discriminator and the results are presented below, along with theoretical curves computed in the method described previously, again allowing for a 0.78° skew angle. In this case the two detectors were transposed when compared with Fig. 8·5 so the D1 dataset has been reduced by a factor of 1.2 for best fit.

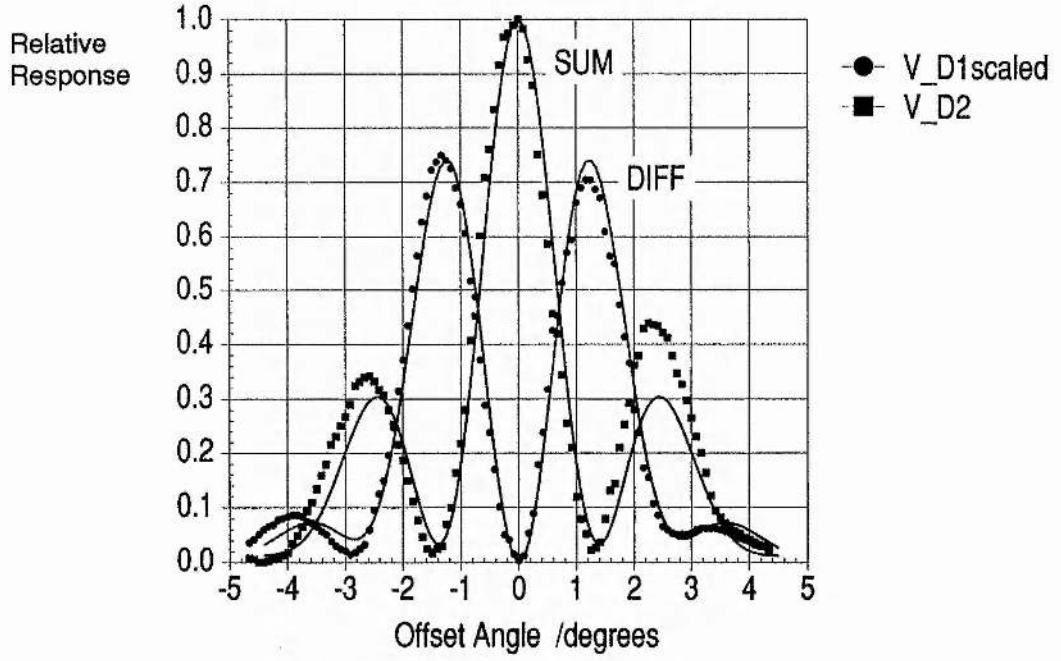


Fig. 8·11 : Complete System Response - Linear Scale

As with the spatial interferometer alone, agreement between theory and experiment is good, particularly within ±2°. The patterns are slightly asymmetric but this is probably due to the slightly

uneven antenna patterns shown in Fig. 8·10. The measured and scaled data are displayed on a log scale below:-

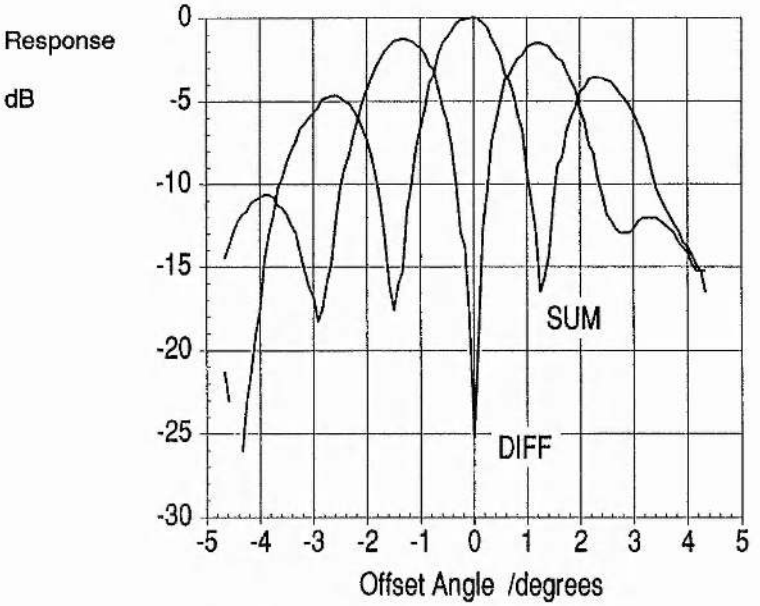


Fig. 8·12 : Complete System Measured Response - Log Scale

The difference channel null is -26dB with respect to the sum channel. The null is shown in detail in the next figure which plots the difference between sum and difference data sets, when expressed in decibels.

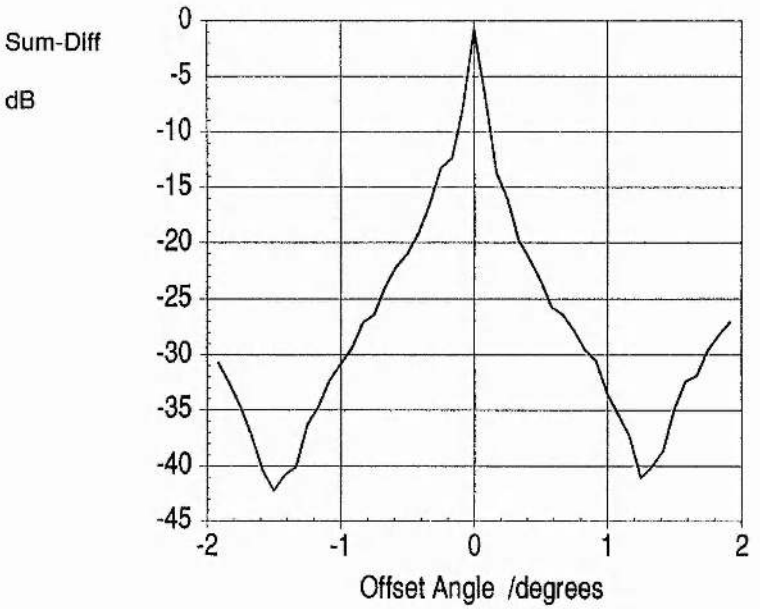


Fig. 8·13 : Detail of Measured Null Rejection

From Fig. 8·13, one can determine the width of the null. At the -12dB level, the null covers slightly less than ±0.2°, which is almost identical to the value obtained for the spatial interferometer alone. This is again less than one tenth of the antenna beamwidth.

Comparing Figs. 8·11, 8·12 & 8·13 with Figs. 8·5, 8·6 & 8·7 show that the combined system operates in a manner very similar to that measured for the spatial interferometer alone. The addition of a pair of frequency discriminators to the spatial interferometer has not changed the angular resolution of the instrument significantly. The theoretical skew angle of 0.78° seems appropriate for both sets of results and is probably due to a small amount of mechanical slackness in the movable roof mirrors of the spatial interferometer. In addition to having different sensitivities, the power detectors are also nonlinear and this may account for some of the distortions in the response curves. A calibrated power meter would be more suitable for making measurements of this type.

8.4.2 Loss in Optics

In order to measure the loss in the receiver optics, a source of known power was coupled into one input port using a feedhorn and lens. The power at the detector ports was measured and referred to the source power. It is known that the insertion loss of a corrugated horn is virtually negligible, but in an optical circuit, lens absorption, truncation and VSWR effects all contribute to an overall loss of signal.

As constructed, the system operates with just a pair of lenses per optical path which is insufficient. This will result in poor coupling and is probably the main contribution to the loss. Disregarding the known loss of -3dB at the first diagonal grid, and a further -3dB because only two detectors are being used, the loss as measured was -2.5dB. Therefore, the total system loss in the receiver is -8.5dB.

In practice, the receiver is operated without lenses on the input ports (which would better define the antenna pattern) so this figure for system loss could easily be nearer -10dB.

The overall gain of the system can be determined more reliably in another way. Measurements taken elsewhere on a duplicate system have shown a total receiver loss of -10dB[5]. Using a source in the far field, the received power is measured first through the receiver optics and then through a calibrated 20dBi feedhorn. The receiver registered 4dB more power than the feedhorn making the system gain 24dBi. Theoretically, the gain of the two-port is 34dBi (for two ±2.25° antenna patterns) so the total loss in the receiver is -10dB, of which -6dB is known to be discarded, so the remaining -4dB represents the optical loss.

An improvement of approximately 2.5dB has been made by adding 45° Faraday rotators after the input grids, thus recovering most of the input power lost at the first diagonal grid. Such rotators offer to recover all of this lost -3dB of signal but have a typical insertion loss of about -0.5dB. The rotators are made from two 1.5mm thick layers of ferrite-loaded flexible plastic sheet[6] sandwiched between quarter wavelength matching layers of unplasticised polyvinylchloride (UPVC)[7]. The layers are glued together with a thin solution of epoxy resin. The ferrite is axially magnetised in the recommended saturation field of 1 Tesla. The rotators performed best around 89GHz and had an insertion loss of -0.3dB but at 94GHz it dropped to approximately -0.8dB. This figure could be improved with more careful manufacture and a better choice of materials[8].

By using two more power detectors/radiometers, another 3dB improvement is possible. The optical loss could also be minimised (to perhaps -1.5dB) with the use of a set of correctly designed lenses. If all these measures were employed, the total system loss could realistically be reduced to about -2dB.

8.5 Conclusion

This chapter has presented a quasi-optical two-port spatial interferometer for which the theoretical and experimental results are in good agreement. The interferometer is capable of accurate sub-beamwidth angular measurements and as such would be suitable for accurately locating millimetre wave beacons.

A complete quasi-optical receiver circuit has been built for the Beacon Detection System which combines a spatial interferometer with two frequency discriminators, as described in Chapter 7. The measured angular resolution of the whole system is almost identical to that of the spatial interferometer alone.

Unfortunately, time did not allow for a full test of the receiver's ability to locate and correlate spread spectrum beacons at St. Andrews. However, a second receiver using a duplicate quasi-optical circuit with dedicated digital signal processing electronics has verified the principle of operation of the complete system[9].

A theoretical analysis of system performance will be presented in Chapter 9.

REFERENCES

[1] Leeson, M.J., *The Application of Quasi-Optical Techniques to Millimetre Wave Radar*, PhD Thesis, St. Andrews, 1992.

[2] Lesurf, J. & Robertson, M.R., "MM-Wave Spatial Interferometry as an Alternative to Radar for Coherent Point Sources", *Intl. Jnl. IR & MM Waves.*, **15** (11), Nov. 1994, pp. 1829-1839.

[3] Skolnik, M., *An Introduction to Radar Systems*, 2nd Ed., McGraw-Hill, 1980.

[4] Leonov, A.I., & Fomichev, K.I., *Monopulse Radar*, Artech House, 1986.

[5] Borrill, J.R., DRA Malvern, Private Communication.

[6] Magnet Developments Ltd., Unit 17, Highworth Industrial Park, Swindon SN6 7NA.

[7] Goodfellow, Cambridge Science Park, Cambridge CB4 4DJ.

[8] Smith, G.M., Unsworth, C.P., Webb, M.R., Lesurf, J.C.G., "Design, analysis and application of high performance permanently magnetised, quasi-optical Faraday rotators", *IEEE MTT-S Digest*, 1994, pp. 293-296.

[9] Borrill, J.R. & Appleby, R., DRA Malvern, Private Communication.

Chapter Nine
Beacon Detection System - Part 3
Analysis and Discussion

9.1 Analysis of Quasi-Optical Circuit

This section presents an analysis of the optical performance of the receiver circuit of the Beacon Detection System which is shown schematically in Fig. 9·1 overleaf. Graphs of theoretical performance are presented which were generated with the aid of computer programs written in 'C' by Jim Lesurf. The techniques for analysing the circuit are quite standard and henceforth this approach could be applied to virtually any quasi-optical circuit.

The analysis concentrates on how the performance of the system varies with the size of the optical elements. Like all the circuits constructed at St. Andrews, the BDS is built using a 'half-cube' method which is flexible for experimental purposes as it is easy to alter. The basic scale factor of such a circuit is the half-cube size and hence the computations have been based around this dimension.

The programs assume that the system has beam paths which are all nominally 12 cubes in length and that there are four lenses equally spaced in each path - note that Fig. 9·1 shows the circuit as implemented which has only two lenses per path. The cubes used have a standard cube size (face width/height/length) of 70mm, mounted on a baseplate with 17.5mm pitch dowel-holes. The lens apertures are 52mm in diameter and the frequency discriminators have arm differences (half-path-length differences) of 20mm.

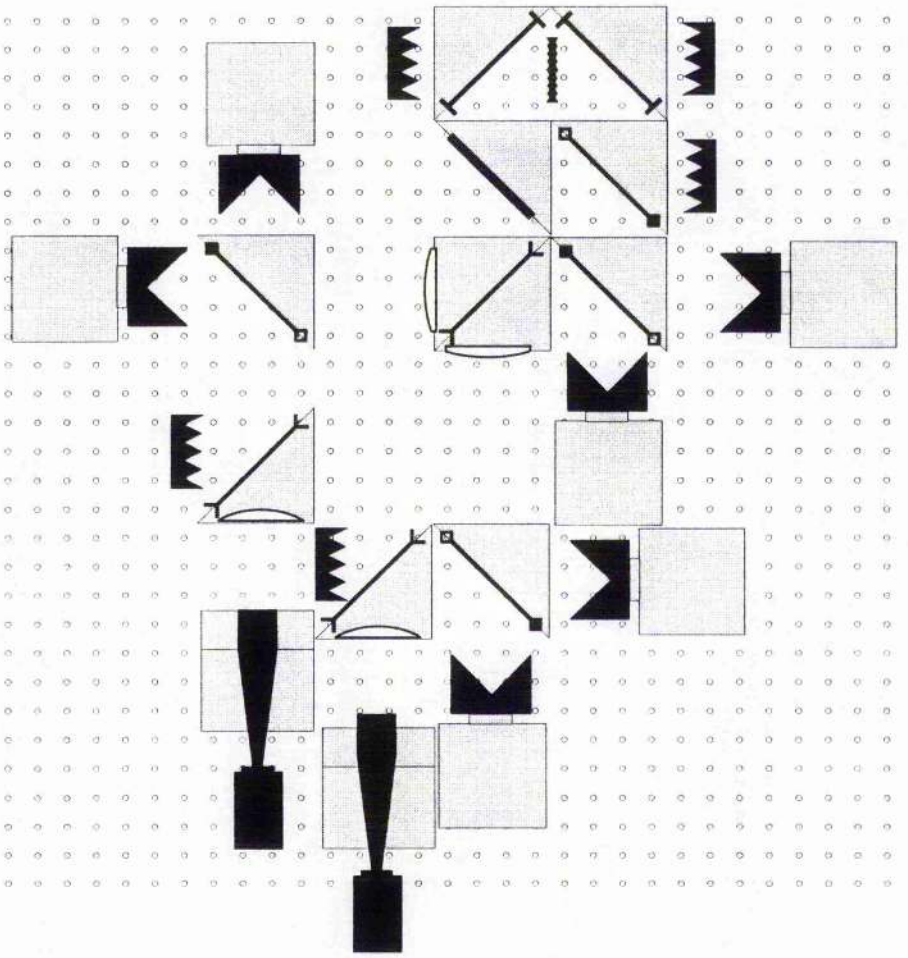


Fig. 9·1 : The BDS QO circuit

9.1.1 Basic Analysis

The overall antenna gain of the receiver is principally determined by two factors. Firstly, the beamwaist size at the input collecting apertures, and secondly, the coupling and propagation losses through the QO circuit itself. For simplicity the analysis assumes the input apertures are the same size as those in the following circuit, but note that this need not be the case. One could arrange to have large collecting apertures and a much more compact processing circuit behind.

As was mentioned in Chapter 1, quasi-optical circuits conventionally make use of fundamental mode Gaussian beams and

they are assumed here. This minimises frequency dependence in circuit performance and makes the analysis far simpler[1]. The use of corrugated scalar feedhorns has also been assumed for the purpose of efficiently launching fundamental mode Gaussian beams. While mirrors would avoid absorption losses, they require a larger, heavier system so it is assumed that lenses are employed for beam focussing.

Hence, the main causes of coupling losses in the system are likely to be lens absorption, truncation at lens apertures and cross-polar coupling losses in the discriminators.

It is worth noting that there are some other minor loss mechanisms which depend on the quality of certain components. Firstly, good free-space wire polarisers have a cross-polar leakage of around −35dB at 90GHz[2] which corresponds to a co-polar loss of less than -0.01dB. Secondly, in a wideband system, lens reflections can produce significant VSWR problems. For a narrowband system the use of blazing and the choice of 'tuned' lens thicknesses can produce VSWR's well below equivalent waveguide systems (i.e. better than 1:1.01)[3]. These two effects will be ignored for the following analysis.

The dielectric losses in a lens depend upon its effective thickness and its loss tangent. Here we assume HDPE lenses which have a refractive index of $\mu = 1{\cdot}53$ and a loss tangent $\delta = 0{\cdot}0003$ at 100 GHz. The attenuation coefficient of the material will be

$$\alpha = 0{\cdot}91\mu\delta f \qquad 9.1$$

where α is in dB/cm, and f is in GHz[4]. For a 'thin' convex lens of focal length, F, the minimum required on-axis thickness, t, will be

$$t = \frac{a^2}{2F(\mu - 1)} \qquad 9.2$$

where a is the radius of the lens aperture. For practical reasons the model assumes that each lens has an extra 2mm thickness to provide a minimum 'lip ring' for holding it in place. In practice, the outer parts of the lens are thinner than at the optical axis so by using the

axial thickness we will overestimate the dielectric losses.

The truncation loss for a fundamental mode beam of size ω will be equivalent to a power coupling (transmission) factor of[5]

$$T = 1 - \mathrm{Exp}\left\{-2\left(\frac{a}{\omega}\right)^2\right\} \qquad 9.3$$

Combining the expressions above we can therefore estimate the losses in coupling through a single lens. Experience with similar systems indicates that one can model a series of lens/apertures by multiplying the individual coupling factors together. Strictly speaking this approach is only approximate since it ignores VSWR effects and the generation of higher-order modes due to diffraction of the truncated beam. However, in most practical cases the truncation loss is small provided we ensure an underillumination ratio, $U = 2a/\omega$, value of around 3, and the lenses are in the near-field of their neighbours (i.e. within $\hat{z} = \lambda z/\pi\omega_0^2 < 1$). The modelling of VSWR and diffraction effects is non-trivial and the more accurately predicted losses are usually smaller than those of the simple model used here.

Cross-polar effects in the frequency discriminators will be dealt with in section 9.1.3.

9.1.2 First Model - Perfect Discriminators

The first 'C' program which was used to model the system performance was written using the assumption that the frequency discriminators produce no coupling or cross-polarisation effects. This situation can be approached at large cube size or by using high-performance refocussing twist-plates instead of roof-mirrors.

Graphs have been generated to show the dependence of various system parameters on cube size. A number of combinations of underillumination ratio, U, and aperture/cube size ratio, d/L, are presented to illustrate the effect they can have. $U = 3.00$ is typical for the lenses designed at St. Andrews and $d/L = 0.740$ represents

the proportions of the actual cubes used in the prototype (52mm apertures in 70mm cubes).

The curves are limited at the lower values of cube size by a condition called "cut-off". This occurs when a lens cannot throw a similar sized beam onto its companions because diffraction effects are too strong. The calculations are halted if cut-off is reached. Note that this is quite different to the cut-off in a waveguide system, below which the wave will not propagate. In this case, propagation continues but the beam is distorted.

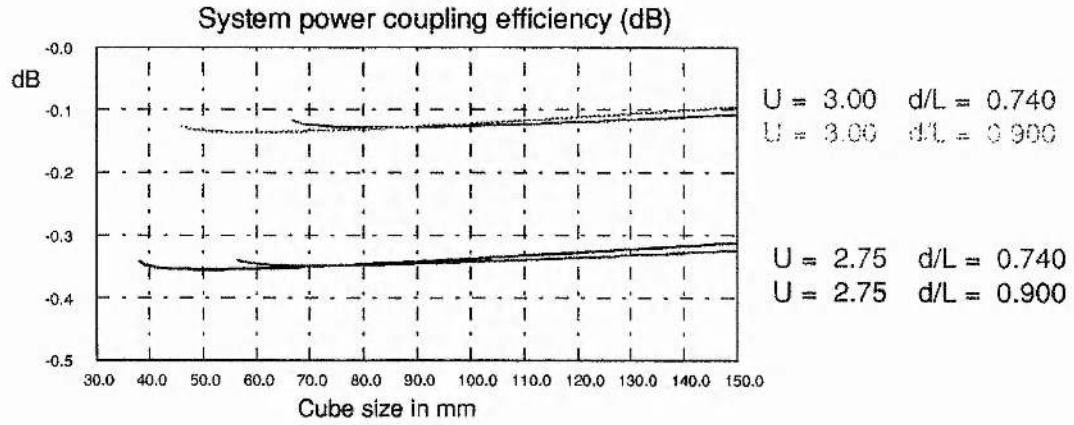


Fig. 9·2 : System power coupling efficiency (perfect discriminators)

Fig. 9·2 illustrates how the power coupling efficiency of the system (the combined effects of dielectric and aperture losses) depends upon cube size for two typical choices of underillumination factor and aperture/cube size ratio. Note that the lower underillumination means the system can operate at lower cube sizes but this is achieved at the expense of slightly greater loss. However, coupling losses of a few tenths of a dB are not serious.

For an underillumination ratio of around 3, the single-port antenna pattern of the system will be very close to Gaussian. The 1/e power half-angle of a single port is shown as a function of cube size in Fig. 9·3, overleaf.

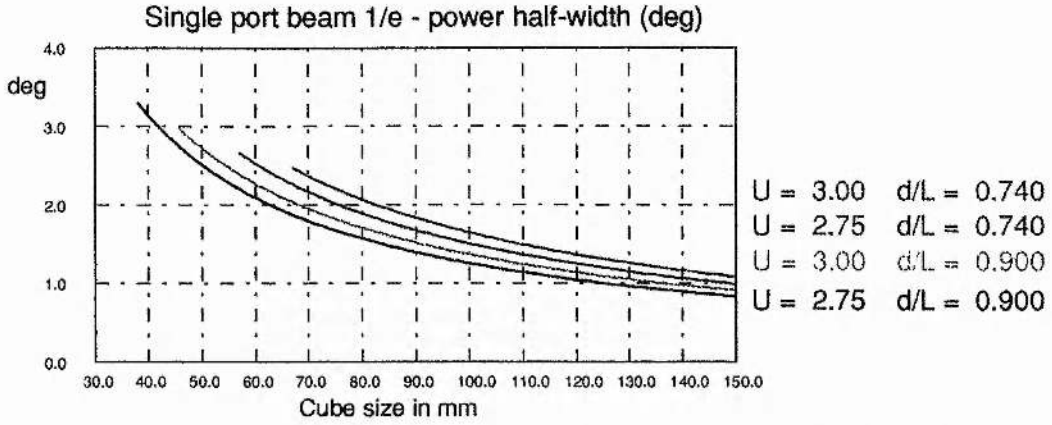


Fig. 9·3 : Single-port output beam 1/e power half-angle

Fig. 9·4 shows how the total on-axis antenna gain of the system depends upon the choice of cube size. The curves displayed are based on the losses and single port beamwidths shown in the previous two graphs. Note that since the system uses a two port spatial interferometer the forward gain is 3dB greater than the power gain of a single port.

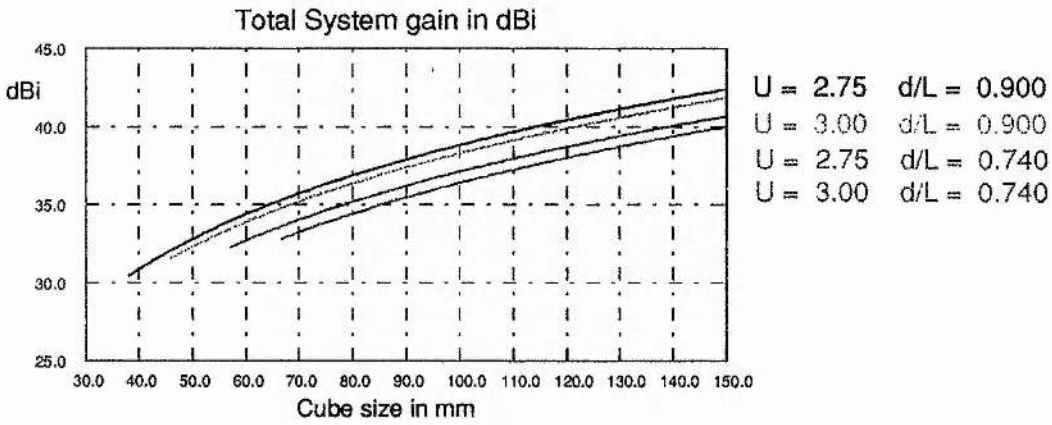


Fig. 9·4 : Total on-axis gain of system

9.1.3 Second Model - Realistic Discriminators

Figs. 9·2-9·4 assume that the frequency discriminators have no significant effect upon performance. A second modified program was written to compute the results of using conventional MPI frequency discriminators with roof-mirrors, as described in Chapter 7.

The main effect of these is to introduce systemic cross-polarisation. This is because diffraction within the MPI causes the recombined beams to have differing widths and phasefront curvatures, since they have propagated different distances after being split apart. These differences mean the beams cannot be perfectly overlaid hence they cannot be made to combine or cancel perfectly. This could be overcome by using a refocussing system such as curved twistplates, for example.

A reduction in cube size means a reduction in beam width, which intensifies the cross-polarisation effect. The effective MPI co-polar power coupling efficiency for a fundamental mode beam is given by[6]

$$M = \sqrt{\left[1 + \left(\frac{A\lambda}{\pi\omega_0^2}\right)^2\right]^{-1}} \qquad 9.4$$

where A is the discriminator arm difference (half the total path difference) and ω_0 is the beamwaist size. The cross-polarisation produced means that an input beam which has power levels, P_v & P_h in the vertical and horizontal planes will produce an emergent beam having powers of

$$P_v' = MP_v + (1 - M)P_h \; ; \; P_h' = MP_h + (1 - M)P_v \qquad 9.5,9.6$$

In the BDS, two discriminators are placed in the spatial interferometer sum and difference beam paths. Here the cross-polarisation produces a slight loss in the discriminator sensitivity. The effect is small, a -20dB cross-polarisation only lowering the sensitivity by 1%.

Fig. 9·5 shows the total system coupling efficiency for a circuit which includes MPI frequency discriminators with a 20mm arm length difference (i.e. a path difference of 40mm). The displayed curves include the effects of lens and truncation losses.

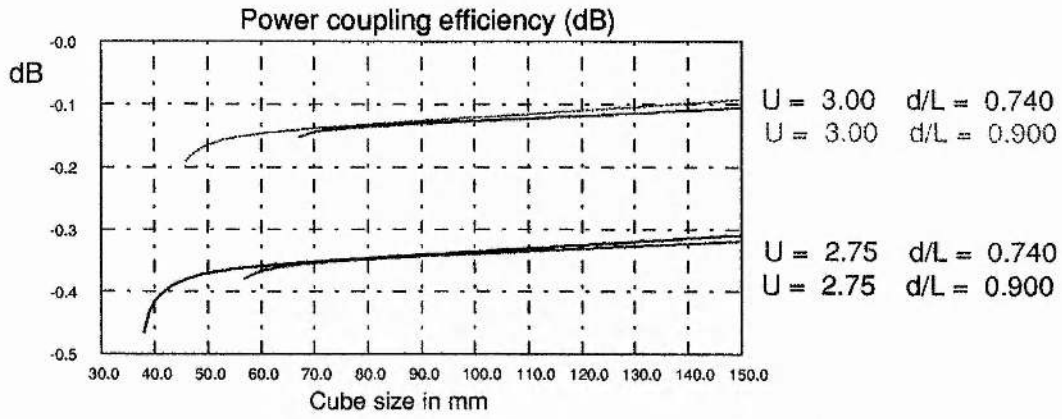


Fig. 9·5 : System co-polar power coupling efficiency

Fig. 9·6 shows the cross-polarisation level produced solely by the discriminators, assuming the polarisers contribute nothing.

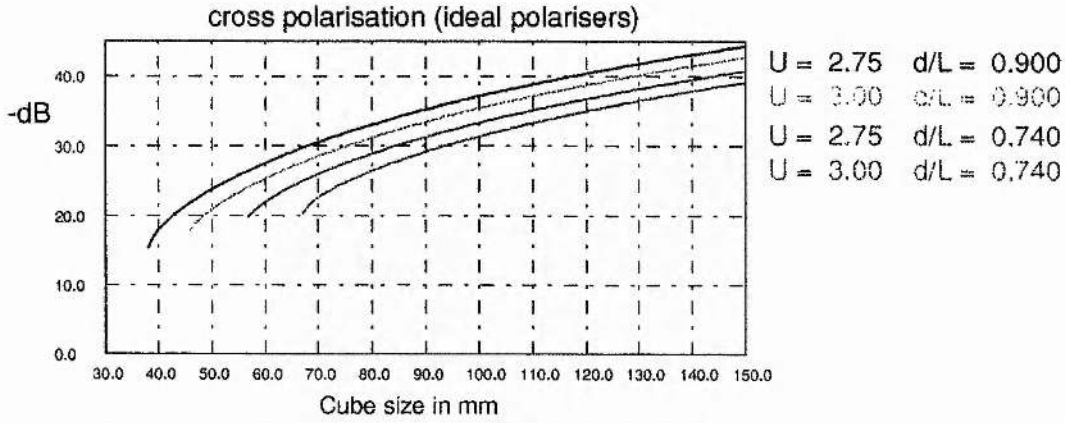


Fig. 9·6 : Cross-polarisation due to MPI frequency discriminators

By looking at Figs. 9·5 and 9·6 and referring back to Fig. 9·4 we can see that the cross-polarisation losses increase rapidly with falling cube size when the size is relatively small. At small cube sizes the MPI cross-polar effect tends to dominate over the lens and truncation losses. Conversely, above a cube size of about 80mm, the MPI effect would become submerged below the general level of polariser imperfections which are likely to limit the cross-polarisation level to the order of -30dB for free standing grids or perhaps -20dB for printed plate polarisers[7].

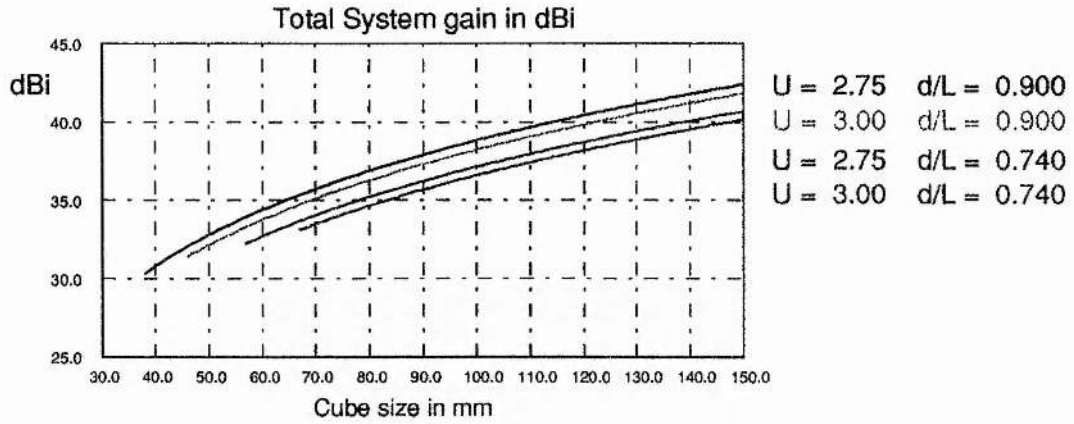


Fig. 9·7 : System on-axis gain with standard MPI discriminators

Fig. 9·7 illustrates the system on-axis total gain including all losses. By comparing this graph with Fig. 9·4 the effect of the frequency discriminators can be seen to be generally small.

Whilst the modular prototype receiver has been constructed on an optical baseplate for experimental flexibility, the system could easily be made more compact. As has been shown, the cube size can be reduced significantly without noticeable degradation of performance, and part of the optical circuit could be "folded up" onto a second layer. To maximise the antenna gain, the input lenses could be made larger than the cube size, provided that they are designed correctly to couple into the following optics.

Free-standing wire grids offer the best performance as polarisers but they are time-consuming and fiddly to construct. Printed plate polarisers are simpler and cheaper to make but have inferior performance. Depending on the application of this system, the printed grids may offer an overall advantage.

All the half cubes, polariser frames and baseplate as used in the modular test set up are made from aluminium. In a final system, circuit flexibility is not an issue and large savings on weight and cost could be made by fabricating the optical elements in a different way, such as with injection moulded plastics.

9.2 Signal-to-Noise Performance

This section deals with the analysis of the signal-to-noise behaviour of the Beacon Detection System receiver. The recovered signal and detected noise levels are considered individually and then combined to produce an expression for the signal-to-noise ratio (SNR) of the whole receiver. A 'C' program was written by Jim Lesurf, based on the results of this section, and used to model the SNR behaviour.

9.2.1 Relative observed AM/FM noise levels

Let us assume that we have an MPI polarising arrangement (a frequency discriminator) feeding a matched pair of power detectors or mixers. Consider initially the system's noise response. The input rf bandwidth to which the detectors can respond is represented by B. The time delay in the MPI is Δ/c where c is the speed of light and Δ is the path difference. The field levels incident on the two detectors at any moment, t, will be

$$E_1 = \tfrac{1}{2}[E\{t\} + E\{t-\Delta/c\}] \;\; ; \;\; E_2 = \tfrac{1}{2}[E\{t\} - E\{t-\Delta/c\}] \quad 9.7,9.8$$

where $E\{t\}$ is the input noise field arriving at any instant via the nominally 'undelayed' path. Common mode delays are irrelevant and can be ignored. If we can arrange that $\Delta/c \geqslant 1/B$, then we can say that the levels $E\{t\}$ and $E\{t - \Delta/c\}$ are statistically independent variables — i.e. one value cannot be predicted from the other. As a result, the field levels falling upon the two detectors are statistically independent. Assigning each detector/mixer a responsivity α we can say that the output voltage at any instant from each will be

$$V_1\{t\} = \alpha E_1^2\{t\} \quad ; \quad V_2\{t\} = \alpha E_2^2\{t\} \qquad 9.9,9.10$$

When demodulating FM signals we use the sum and difference values

$$V_{FM} = V_1 - V_2 \quad ; \quad V_{AM} = V_1 + V_2 \qquad 9.11,9.12$$

Note that in Chapter 7, the discriminator response is given by difference-divided-by-sum (eqn. 7.22) so that fluctuations in signal

power level are suppressed. However, here we are considering only the input noise, so there is no signal power, and the simpler expression of eqn. 9.11 is sufficient.

From the basic statistical independence and the system's symmetry we can conclude that the output noise power appears equally divided between the observed levels, V_{FM} and V_{AM}. In simple terms we can therefore expect half the input noise fluctuations to appear as demodulated AM measurement noise and half as FM measurement noise.

The noise contributions generated inside the detector/mixers and their associated amplifiers are also statistically independent, hence a similar argument applies to mixer/amplifier noise up to the point where the detector output signals are combined to recover AM/FM measurement information.

Although a system employing a pair of mixers plus wideband IF amps and crystal detectors is more complex, its behaviour is, in information terms, equivalent provided the bandwidth is large enough to satisfy the requirement indicated above. In this sense, radiometers are just very sensitive power detectors. Note however, that any noise sidebands on the LO power used to pump the mixers will generate common mode effects. When used as wideband power radiometers this has no first order effect on the FM noise level observed since one isn't recovering any information regarding the relative frequency of the LO. It may, however, increase the observed AM noise level since any modulation of the LO level (LO AM sidebands) may modulate the responsivity (i.e. the conversion gain) of the mixers. Note that some of this may be converted to apparent FM noise if the mixers are not well matched. This is because it will then not be suppressed from the FM observation by taking the difference between the two detected power levels.

9.2.2 The observed AM/FM noise level

The following discussion defines the receiver as being a single sideband (SSB) receiver. Reworking the argument in terms of a double sideband (DSB) receiver is discussed in a later section but it should be emphasised that the final results are equivalent.

Suppose the field of view has an effective input brightness temperature, T_a, and that each mixer/amplifier has an effective noise temperature, T_{nSSB}, and a conversion/amplifier total gain of G, up to the crystal detectors which have responsivity, η. The noise power generated mean d.c. levels from the crystal detectors will be

$$V_1 = \eta G B_{IF} k\left(T_{nSSB} + T_a/2\right) \; ; \; V_2 = \eta G B_{IF} k\left(T_{nSSB} + T_a/2\right) \quad 9.13,9.14$$

where B_{IF} is the IF amp bandwidth and k is Boltzmann's constant.

From the above we can define a total input noise power temperature level $T' = T_{nSSB} + T_a/2$. From radiometer theory, a system which has a post-detection time constant, τ_x, will output an rms noise fluctuation level equivalent to a temperature uncertainty of

$$\Delta T \; = \; \frac{T'}{\sqrt{B_{IF}\tau_x}} \qquad 9.15$$

From the arguments of statistical independence and symmetry outlined earlier, the output rms voltage noise level will then be

$$\Delta V \; = \; \eta G T' k \sqrt{\frac{B_{IF}}{\tau_x}} \qquad 9.16$$

This represents the likely voltage error for both amplitude (sum) and frequency (difference) measurements averaged over the time, τ_x. In many practical cases the mixer/detector noise level is significantly larger than the field of view temperature in which case we can assume that $T' \approx T_{nSSB}$ and the output rms noise level is

$$\Delta V \; \approx \; \eta G T_{nSSB} k \sqrt{\frac{B_{IF}}{\tau_x}} \qquad 9.17$$

unaffected by thermal sources in the field of view.

9.2.3 The magnitude of the recovered FM signal

Consider now the situation when the system is receiving a coherent signal whose frequency at any instant is

$$f\{t\} = f_0 + \delta f\{t\} \qquad 9.18$$

and whose average power is steady at the level, P. (i.e. no AM) Then we can say f_0 is the nominally unmodulated 'carrier' frequency and $\delta f\{t\}$ represents the FM modulation we wish to recover. The modulation is assumed to have an rms magnitude Δf.

The signal is demodulated with an MPI frequency discriminator such that

$$V_{fm}\{t\} = \eta GP \sin\left\{\frac{2\pi D.\delta f\{t\}}{c}\right\} \qquad 9.19$$

where the path difference, D, is chosen to give an output of zero for an input frequency of f_0. Provided that $\delta f < c/4D$ this can be approximated to

$$V_{fm}\{t\} = \eta GP\frac{2\pi D.\delta f\{t\}}{c} \qquad 9.20$$

Therefore the rms magnitude of the demodulated FM signal pattern will be approximately ($\langle\,\rangle$ indicate an average quantity.)

$$\sqrt{\langle V^2_{fm}\rangle} = \eta GP\frac{2\pi D.\Delta f}{c} \qquad 9.21$$

The detailed analysis of discriminator behaviour in Chapter 7 indicates that satisfactory behaviour can be obtained when $D.\Delta f/c \approx 1/12$. Using this value one obtains the expression

$$\sqrt{\langle V^2_{fm}\rangle} = \frac{\pi}{6}\eta GP \approx \frac{1}{2}\eta GP \qquad 9.22$$

The values chosen here are not precise for all cases, but it should be possible to ensure that $V_{fm} \approx \frac{1}{2}\eta GP$ in most situations of practical interest.

The correlation process used to recognise the demodulated signal can be treated as a two-stage process. The recovered signal pattern, $V_{fm}\{t\}$ is first multiplied by a function, $p'\{t\}$, which is a

normalized (i.e. unity rms magnitude) version of the pattern, $p\{t\}$, which we are seeking. The multiplied output is then integrated over the chosen measurement time, T. This process produces an output

$$V_0 = \frac{1}{\tau}\int_0^T V_{fm}\{t\}.p'\{t\}\,dt \qquad 9.23$$

where τ is the integrator's time constant value ($\tau = RC$ for a normal analog integrator).

Since the time averaged value of $\langle p\{t\}.p\{t\}\rangle = \langle p^2\{t\}\rangle$ it follows that, when we have matched the correct pattern (i.e. when $V_{fm}\{t\}$ matches $p'\{t\}$), that

$$\langle V_{fm}\{t\}.p'\{t\}\rangle = \sqrt{\langle V^2{}_{fm}\rangle} \qquad 9.24$$

i.e. the multiplication process recovers the rms level of the input signal. Hence we can say that the final measured signal output from the correlation integration will be

$$V_0 \approx \frac{\eta GPT}{2\tau} \qquad 9.25$$

9.2.4 The Noise Output from the Correlator

Consider now two noise input voltages entering the correlation system. Let us consider how the multiplication section of the correlator affects the noise properties of these two inputs.

Suppose that each of the two noise sources input to the multiplier are 'white' and extend over a bandwidth, $B = 1/2\tau_x$, which covers f_{min} to f_{max}, limited by the post-detection time constant.

This bandwidth can be sub-divided into $(I + 1)$ frequency components $f_i = f_{min} + i.\delta f$ where $\delta f = B/(I + 1)$. Then we can say that the noise in this band can be represented in terms of a series of frequency components

$$v_i\{t\} = \sum_{i=o}^{I} a_i \sin\{2\pi f_i t + \phi_i\} \quad ; \quad v_j\{t\} = \sum_{j=o}^{I} b_j \sin\{2\pi f_j t + \phi_j\} \qquad 9.26,9.27$$

where a_i, b_i and ϕ_i,ϕ_j are the amplitudes and relative phases of the components. Fig. 9·8 shows the power spectrum of the noise.

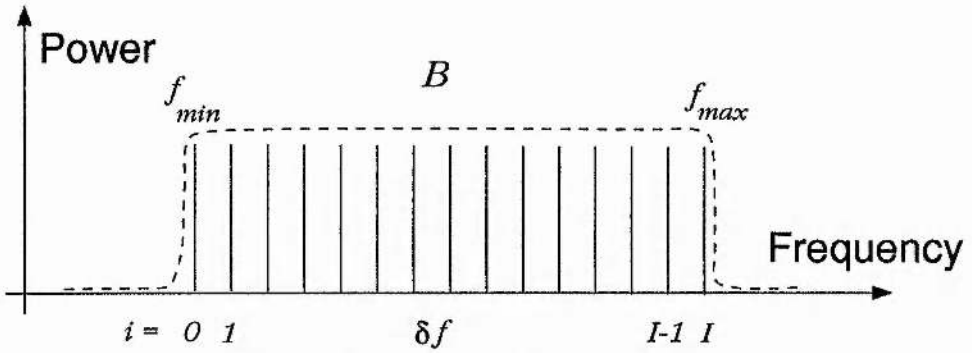


Fig. 9·8 : White noise spectrum input to correlator

For white noise, all the amplitudes can be taken as equal to a (or b), a suitable mean value. Therefore, each frequency component contributes a time averaged power of $\frac{a^2}{2}$ (or $\frac{b^2}{2}$) to the total input signal power level. Since these are spaced δf apart, the signals have effective noise power spectral densites (NPSDs) of

$$S_{in(a)} = \frac{a^2}{2\delta f} \text{ and } S_{in(b)} = \frac{b^2}{2\delta f} \qquad 9.29,9.30$$

(note these are constants because the noises are 'white') and the input powers can be defined as

$$P_{in(a)} = \langle v_i^2\{t\}\rangle = S_{in(a)}.B = \frac{a^2B}{2\delta f} = (I+1)\frac{a^2}{2} \qquad 9.31$$

$$P_{in(b)} = (I+1)\frac{b^2}{2} \qquad 9.32$$

So the <u>total</u> input power is

$$P_{in} = (I+1)\frac{(a^2+b^2)}{2} \qquad 9.33$$

Applying the two noise waveforms, $v_i\{t\}$ and $v_j\{t\}$ to a multiplier of gain K, we have

$$v_m\{t\} = K.v_i\{t\}.v_j\{t\} \qquad 9.34$$

or

$$v_m\{t\} = K \sum_{i=o}^{I} a \sin\{2\pi f_i t + \phi_i\} \times \sum_{j=o}^{I} b \sin\{2\pi f_j t + \phi_j\} \qquad 9.35$$

from which we get

$$v_m\{t\} = Kab \sum_{i=o}^{I} \sum_{j=o}^{I} \left[\sin\{2\pi f_i t + \phi_i\} \times \sin\{2\pi f_j t + \phi_j\}\right] \qquad 9.36$$

Now using

$$2 \sin P \sin Q = \cos(P - Q) - \cos(P + Q) \qquad 9.37$$

we get

$$v_m\{t\} = K\frac{ab}{2} \sum_{i=o}^{I} \sum_{j=o}^{I} \left[\cos\{2\pi(f_i - f_j)t + \Theta_{ij}\} - \cos\{2\pi(f_i + f_j)t + \Omega_{ij}\}\right] \qquad 9.38$$

where

$$\Theta_{ij} = \phi_i - \phi_j \quad \text{and} \quad \Omega_{ij} = \phi_i + \phi_j \qquad 9.39,9.40$$

The expression for the frequency offsets $f_i \pm f_j$ can be replaced with $(i \pm j)\delta f$ to yield

$$v_m\{t\} = K\frac{ab}{2} \sum_{i=o}^{I} \sum_{j=o}^{I} \left[\cos\{2\pi(i - j)\delta f t + \Theta_{ij}\} - \cos\{2\pi(i + j)\delta f t + \Omega_{ij}\}\right] \qquad 9.41$$

By considering the individual terms in the above series we can now discover the overall spectral shape of the output from the multiplier when the inputs have uniform spectral density over the bandwidth B.

There are three cases which must be dealt with separately - the terms contributing to the dc level, the remaining difference terms and the sum terms.

Firstly, for the case of $i - j = 0$, there are $(I + 1)$ contributions at dc from the difference term which give rise to a mean output level voltage of

$$\langle v_{dc} \rangle \approx K(I + 1)\frac{ab}{2} \qquad 9.42$$

Secondly, in the remaining difference terms where $i \neq j$ there

is just one pair of components where $|i - j| = I$ each of which produces output at the frequency $f = I.\delta f$ (N.B. one of these is at $-f$ but $\cos(x) = \cos(-x)$ so the negative sign is not observable.), two pairs of components where $|i - j| = (I - 1)$ producing contributions at the frequency $f = (I - 1).\delta f$, three pairs where $|i - j| = (I - 2)$ at $f = (I - 2).\delta f$, etc.

In general, the number of contributions from the <u>difference</u> term at a frequency

$$f_q = (I - q).\delta f = \left(\frac{B}{\delta f} - 1 - q\right).\delta f \qquad 9.43$$

will be

$$M_q = 2(q + 1) \qquad 9.44$$

provided that $0 \leqslant q < I$.

This gives rise to a triangular noise spectrum which decreases with frequency from dc to B - e.g. Fig. 9·9 where $I = 9$.

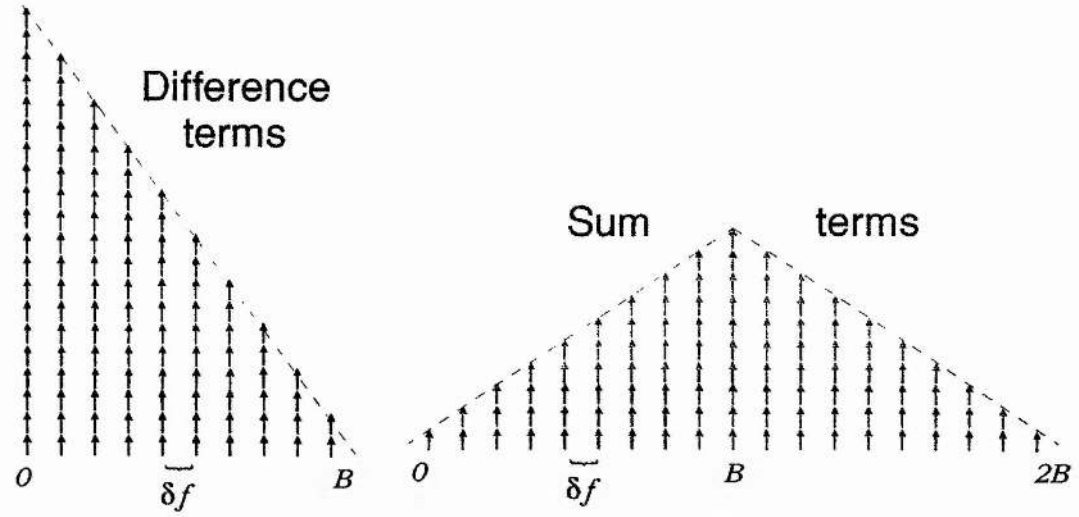


Fig. 9·9 : Separate noise spectra of difference and sum terms

Thirdly, there are the sum terms. There will be one contribution for $|1 + j| = 0$ at frequency $f = 0$, two contributions for $|1 + j| = 1$ at $f = \delta f$, three contributions for $|1 + j| = 2$ at $f = 2\delta f$, rising to I contributions for $|1 + j| = I$ at $f = I\delta f$. Then the trend decreases to $I - 1$ contributions for $|1 + j| = I + 1$ at $f = (I + 1)\delta f$,

down to one contribution for $|1 + j| = 2I$ at frequency $f = 21\delta f$. This gives rise to a triangular peak in the noise spectrum as shown in Fig. 9·9, which rises from dc to a peak at B and decreases again to zero at $2B$. The height of the peak is half that of the difference contributions.

When the sum and difference terms are taken together, the result is a triangular noise spectrum decreasing with frequency from dc to $2B$ - see Fig. 9·10.

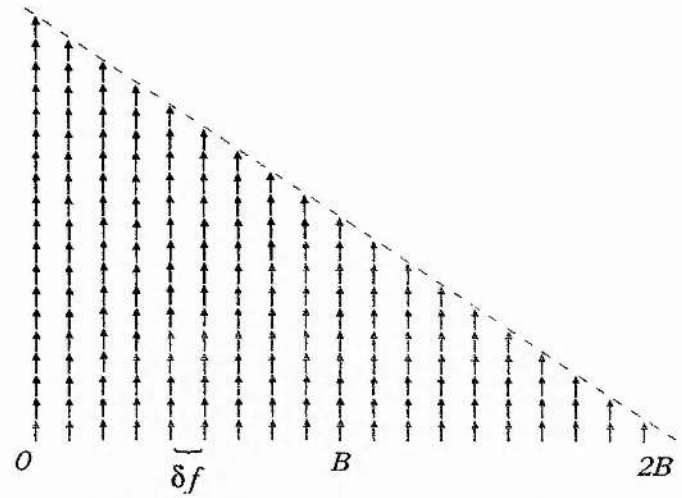


Fig. 9·10 : Overall noise spectrum at multiplier output

The envelope of this spectrum is of the form $y = mx + c$ where the gradient m, is negative and c is the y-axis intercept. The equation of this line must now be determined.

In general, the number of terms from the combined output at a frequency

$$f_p = (2I - p).\delta f = \left(\frac{2B}{\delta f} - 2 - p\right).\delta f \qquad 9.45$$

will be

$$M_p = p \qquad 9.46$$

provided that $0 \leqslant p < 2I$.

For any specific signal, the output power and voltage produced at each frequency will depend on the phase relationship

between the terms. However, for a typical unpredictable signal, the phase relationships will be apparently random. Each term taken individually will produce fluctuations whose mean power level is $\frac{1}{2}\left[K\frac{ab}{2}\right]^2$. Since the phases of the various pairs of contributions are uncorrelated, the overall power of the fluctuations they produce at frequency, f_p, will typically be

$$P_p \approx M_p.\frac{1}{2}\left[K\frac{ab}{2}\right]^2 = \frac{p}{2}\left[K\frac{ab}{2}\right]^2 \qquad 9.47$$

The effective noise power spectral density this produces will therefore be

$$S_{out}\{f\} = \frac{P_p}{\delta f} = \frac{p}{2\delta f}\left[K\frac{ab}{2}\right]^2 = \frac{1}{\delta f}\left(\frac{B}{\delta f} - 1 - \frac{f_p}{2\delta f}\right)\left[K\frac{ab}{2}\right]^2 \qquad 9.48$$

Since $\frac{B}{\delta f} >> 1$ this expression can be simplified to

$$S_{out}\{f\} \approx \frac{1}{\delta f}\left(\frac{B}{\delta f} - \frac{f_p}{2\delta f}\right)\left[K\frac{ab}{2}\right]^2 = K^2\left(B - \frac{f_p}{2}\right)\left[\frac{a^2}{2\delta f}\right]\left[\frac{b^2}{2\delta f}\right] \qquad 9.49$$

$$\therefore \quad S_{out}\{f\} \approx BK^2\left(1 - \frac{f_p}{2B}\right)S_{in(a)}.S_{in(b)} \qquad 9.50$$

Now in the correlator, one input noise spectrum is normalised, so we have

$$S_{in(b)} = \frac{b^2}{2\delta f} = \frac{W}{B} \qquad 9.51$$

where W is a constant power. If we substitute this into the previous equation we get

$$S_{out}\{f\} \approx W.K^2\left(1 - \frac{f_p}{2B}\right)S_{in(a)} \qquad 9.52$$

Replace $B = \frac{1}{2\tau_x}$, and f_p with the continuous variable f and we finally have a noise power spectral density (in units of Volts^2/Hz) of

$$S_{out}\{f\} \approx W.K^2(1 - \tau_x f)S_{in(a)} \qquad 9.53$$

This expression describes a noise spectral density of the form as shown in Fig. 9·10. It shows that the output NPSD is related to the input NPSD by a $-f$ relationship and scaled by the multiplier gain and the power of the normalised noise spectrum.

This equation can now be inserted into the equation for the noise output of an integrator

$$N = \int_0^\infty G_i\{f\} \, .S_o\{f\} \, df \qquad 9.54$$

where the effective power gain (in volts^2 rather than watts) of an analog integrator is[8]

$$G_i\{f\} = \frac{\sin^2(\pi f T)}{(\pi f \tau)^2} \qquad 9.55$$

Therefore we can say

$$N = \int_0^\infty \frac{WK^2(1 - \tau_x f)\, 2B_{IF}(\eta G T_{nSSB} k)^2 . \sin^2(\pi f T)}{(\pi f \tau)^2} df$$

$$= 2WK^2 B_{IF}(\eta G T_{nSSB} k)^2 \int_0^\infty \frac{(1 - \tau_x f) . \sin^2(\pi f T)}{(\pi f \tau)^2} df \qquad 9.56$$

Now the time constant of the crystal detector τ_x is very much smaller than the time constant of the integrator τ, so τ_x can be neglected and we can set $(1 - \tau_x f) = 1$. This leaves a standard integral[8] and we are left with

$$N = 2WK^2 B_{IF}(\eta G T_{nSSB} k)^2 \int_0^\infty \frac{\sin^2(\pi f T)}{(\pi f \tau)^2} df$$

$$= 2WK^2 B_{IF}(\eta G T_{nSSB} k)^2 . \frac{T}{2\tau^2} \qquad 9.57$$

Therefore, the output noise power of the integrator is

$$N = WB_{IF} T \left(\frac{K \eta G T_{nSSB} k}{\tau} \right)^2 \qquad 9.58$$

It should be noted that the normalised power in the reference noise input, W, and the multiplier gain, K, both have unity value. However, they are retained in these expressions to maintain correct dimensionality.

9.2.5 The S/N Ratio of the FM measurement

The final observation S/N power ratio will therefore be

$$\frac{S}{N} = \left(\frac{V_0}{V_n}\right)^2 = \frac{T}{WB_{IF}} \cdot \left(\frac{P}{2KkT_{nSSB}}\right)^2 \qquad 9.59$$

From this final expression we can see that the system's behaviour should not depend upon the details of gain values, time constants, etc, provided that these are consistent with the general assumptions stated. By rearranging the above expression we can say that the signal power required for a given observed signal-to-noise ratio will be (N.B. $W = K = 1$)

$$P = 2KkT_{nSSB}\sqrt{\frac{WB_{IF}}{T} \cdot \frac{S}{N}} \qquad 9.60$$

For some purposes it is interesting to compare this required signal power with the input receiver noise level, $P_n = kT_{nSSB}B_{IF}$. This leads to the result

$$\frac{P}{P_n} = 2\sqrt{\frac{1}{B_{IF}T} \cdot \frac{S}{N}} \qquad 9.61$$

Fig. 9·11 shows how the required signal power depends upon the FM system bandwidth on the basis of some assumptions typical for a system based on the use of heterodyne radiometers.

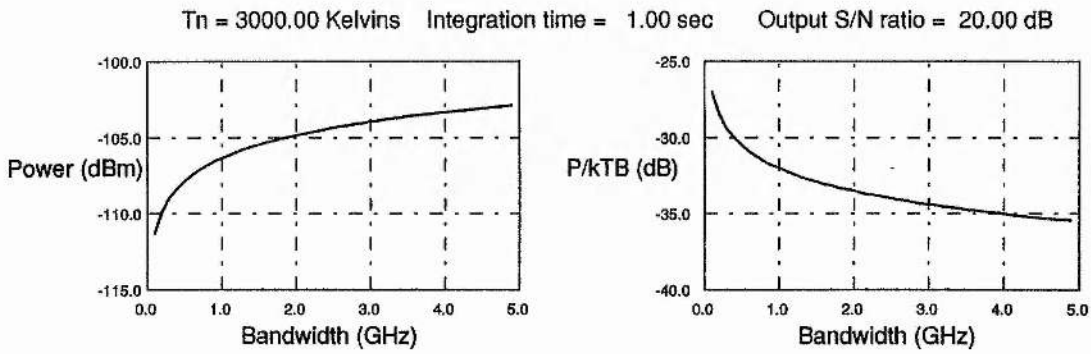


Fig. 9·11 : Required signal power for a typical wideband FM system

9.2.6 Extension for DSB receivers

If we are to define the receiver behaviour as DSB, the effective input bandwidth is $2B_{IF}$ and the DSB noise temperature is $\frac{1}{2}T_{nSSB}$. The temperature uncertainty is still defined as in eqn. 9.15 but since the noise temperature is half that for SSB operation, ΔT is also halved. Now the minimum required input power is $P = k\Delta T_{DSB}2B_{IF}$ for DSB and $P = k\Delta T_{SSB}B_{IF}$ for SSB, so the powers are equal since the doubling in bandwidth is offset by the halving in noise temperature.

Eqn. 9.60 derives the required input signal power for a given signal-to-noise ratio, IF bandwidth, integration time and SSB noise temperature. For a DSB derivation, the noise temperature is halved and there is an additional factor of two on the right hand side due to the additional sideband, but the overall result is the same as quoted.

9.2.7 Note on IF Bandwidths

Note that the effective bandwidth of a radiometer is not simply the -3dB bandwidth quoted for the IF amplifiers. In practice, it can be quite a bit larger, of the order of 1.5× greater or more, depending on the order of filter used to roll-off the amplifier response. It is this effective bandwidth which should be used for B_{IF} in the above.

Additionally, this analysis assumes that the signal transmission occupies the whole of the receiver bandwidth. Ideally with a DSB system we would choose $B_{IF} = \Delta f$. A smaller IF would lose some signal. A wider bandwidth presents an increased noise level to the crystal, hence demanding a higher signal power to obtain the required S/N ratio. In practice therefore, we can say that the minimum required signal power, P_{min}, will occur when $B_{IF} = \Delta f$. A wider IF bandwidth will necessitate a signal power of

$$P = P_{min}\sqrt{\frac{B_{IF}}{\Delta f}} \qquad 9.62$$

For example, assuming a DSB receiver with $B_{IF} = 6$ GHz and a signal having $\Delta f = 500$ MHz, then the required signal power is $\sqrt{12}$ times bigger than it need be. As an alternative way of looking at this, for a given signal power, the S/N ratio will be $\frac{1}{12}$ of what it could be.

Note that with a DSB receiver for which the signal does cover the full DSB bandwidth, the IF amps are AC coupled so some signal will be lost around DC. Provided the low-frequency roll-off of the amps are indeed low, this effect may not be serious. It may also depend on the frequency distribution of the modulated carrier, which is a function of the characteristics of the modulating waveform.

9.3 Probabilities of Detection and False Alarm

In order for the Beacon Detection System to operate without error, one must determine a minimum signal to noise ratio at the output of the correlator which then dictates the required source power of the beacon for a given set of operating conditions. The minimum output SNR is determined through a statistical assessment of the noise properties of the correlator output.

Correlating a pseudo-random waveform with a stored copy of itself produces a correlation peak above a baseline of pseudo-noise. The correct detection of this correlation peak is akin to the problem of correctly detecting a dc level in Gaussian noise. A threshold level is set to trigger the condition of "peak exists - correlation detected". This must be set to identify correctly as many real correlations as possible whilst minimising the number of false alarms caused by noise exceeding the threshold. This can be represented in the following probability diagram.

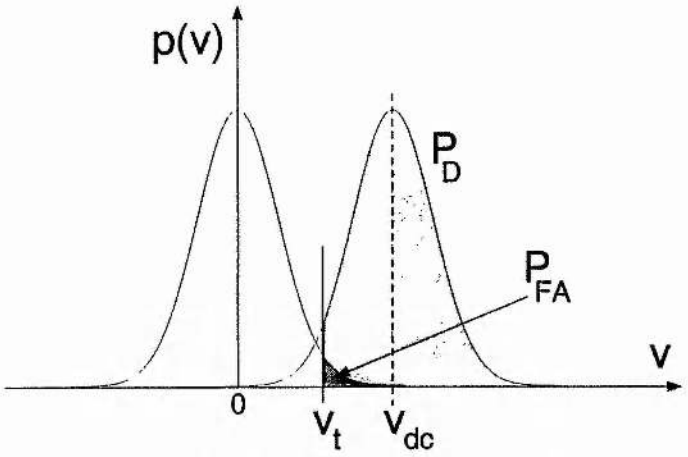


Fig. 9·12: Threshold Detection in Gaussian Voltage Noise

The left-hand Gaussian represents the zero-mean noise and the right-hand Gaussian depicts the dc level, v_{dc}, with Gaussian noise superimposed upon it. The threshold level, v_t, is shown in the upper tail of the noise distribution. The dark-shaded area is the probability that a noise spike will exceed the threshold - this is the probability of false alarm, P_{FA}. The light-shaded area is the probability that a real (i.e. noisy) correlation peak exceeds the threshold - the probability of detection, P_D.

The required values of P_{FA} and P_D are determined by the operational requirements of a given application and henceforth will not be discussed here. Examples will be used where appropriate.

It is possible to compute these probabilities for a range of signal to noise values (range of v_{dc}) and this is done in the following way. Let us assume that the waveforms input to the correlator are Gaussian in probability distribution and white in power spectral density (PSD). The pseudo-noise baseline now has a triangular PSD (see section 9.2.4) but its probability distribution is still Gaussian. In general, a Gaussian voltage probability distribution has the form

$$p(v) = \frac{1}{\sqrt{2\pi}.\sigma} \exp\left(-\frac{(v - v_o)^2}{2\sigma^2}\right) \qquad 9.63$$

where v_o is the mean value, σ is the variance and σ^2 is the standard deviation. For a voltage distribution such as this, the variance is

equivalent to the ac rms voltage, i.e. $v_{rms(ac)} = \sigma$.

Now the single test probability that a given voltage will be greater than a certain threshold, v_t, is

$$P = \frac{1}{\sqrt{2\pi}.\sigma} \int_{v_t}^{\infty} \exp\left(-\left(\frac{v - v_o}{\sqrt{2}.\sigma}\right)^2\right) dv \qquad 9.64$$

If we substitute

$$z = \frac{v - v_o}{\sqrt{2}.\sigma} \quad \Rightarrow \quad v = v_o + \sigma\sqrt{2}.z \quad \Rightarrow \quad dv = \sigma\sqrt{2}.dz \qquad 9.65$$

we get

$$\begin{aligned} P &= \frac{1}{2} \cdot \frac{2}{\sqrt{\pi}} \int_{\frac{v_t - v_o}{\sqrt{2}.\sigma}}^{\infty} \exp\left(-z^2\right) dz \\ &= \frac{1}{2} \cdot \operatorname{Erfc}\left(\frac{v_t - v_o}{\sqrt{2}.\sigma}\right) \\ &= \frac{1}{2} \cdot \left[1 - \operatorname{Erf}\left(\frac{v_t - v_o}{\sqrt{2}.\sigma}\right)\right] \end{aligned} \qquad 9.66$$

where $\operatorname{Erf}(x)$ and $\operatorname{Erfc}(x)$ are the error function and complementary error function respectively[9] and are defined for $x > 0$, which implies that the last equation is only valid for $v_t > v_o$. This condition is satisfactory for the case of P_{FA} because $v_o = 0$ and v_t is always positive. However, for the case of P_D the expression needs to be modified to cover the possibility, as illustrated, that $v_t < v_o$. Computationally, $\operatorname{Erf}(x)$ is replaced with $\operatorname{Sgn}(x).\operatorname{Erf}(\operatorname{Abs}(x))$, where $\operatorname{Sgn}(x)$ is the sign of x, $\operatorname{Abs}(x)$ is the absolute value of x, and x lies in the range $[-\infty, \infty]$.

It is also possible to express the probability in terms of power signal to noise ratio (S/N) by substituting $S = v^2$ and $N = \sigma^2$.

The error function is calculated using the following polynomial approximation[10]

$$\operatorname{Erf}(z) = 1 - \left[at + bt^2 + ct^3\right] \exp\left(-z^2\right) \qquad 9.67$$

where

$$t = \frac{1}{1 + Pz}$$

$$P = 0.47047;\ a = 0.3480242;\ b = -0.0958798;\ c = 0.7478556$$

I have written two BASIC programs (listed in Appendix 3) to generate curves of P_D versus *SNR* for various levels of P_{FA} and the results are presented below. I used the voltage form of the equations for simplicity and converted to power *SNR* before plotting. In order to calculate the data sets over a range of *SNRs*, I set the noise voltage equal to unity and varied the signal voltage. In this way, the calculation for P_{FA} becomes normalised (because $v_o = 0$ and $\sigma = 1$) so a given value of threshold will set a certain probability of false alarm. The values of P_{FA} were easily calculated using

$$P_{FA} = \frac{1}{2} \cdot \left[1 - \mathrm{Erf}\left(\frac{v_t}{\sqrt{2}}\right)\right] \qquad 9.68$$

The results are presented in Fig. 9·13 versus power threshold to noise ratio, *TNR* (*TNR* = $10 \log(v_t^2)$ in this case). Note that this depicts the probability of false alarm for a single test. For datasets with many points contributing to the false alarm probability, the expression must be modified. This is dealt with later.

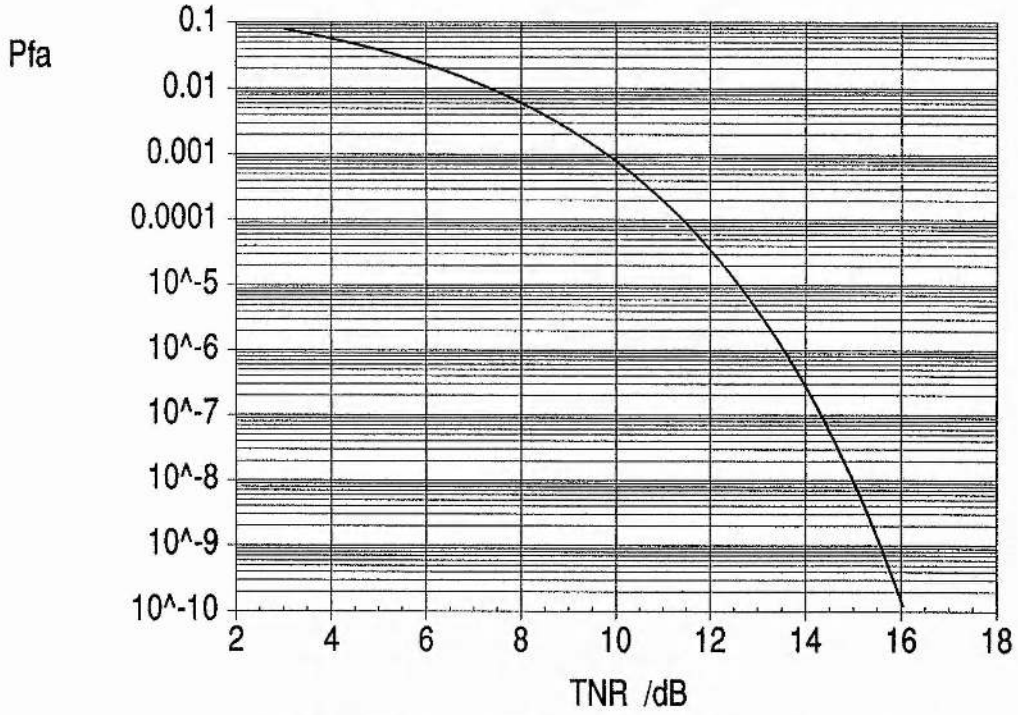


Fig. 9·13 : Probability of False Alarm for a Single Test

From this graph I was able to pick the required values of *TNR*

which set a given P_{FA} and use them to calculate the probability of detection. The equation for this is

$$P_D = \frac{1}{2} \cdot \left[1 - \mathrm{Erf}\left(\frac{v_t - v_s}{\sqrt{2}}\right)\right] \qquad 9.69$$

where the threshold v_t is determined from Fig. 9·13 ($v_t = \sqrt{TNR}$) and v_s is the voltage signal to noise level ($\sigma = 1$ again). The power SNR is then calculated as $SNR = v_s^2$. The results are shown in Fig. 9·14 for values of P_{FA} ranging from 10^{-2} to 10^{-9}.

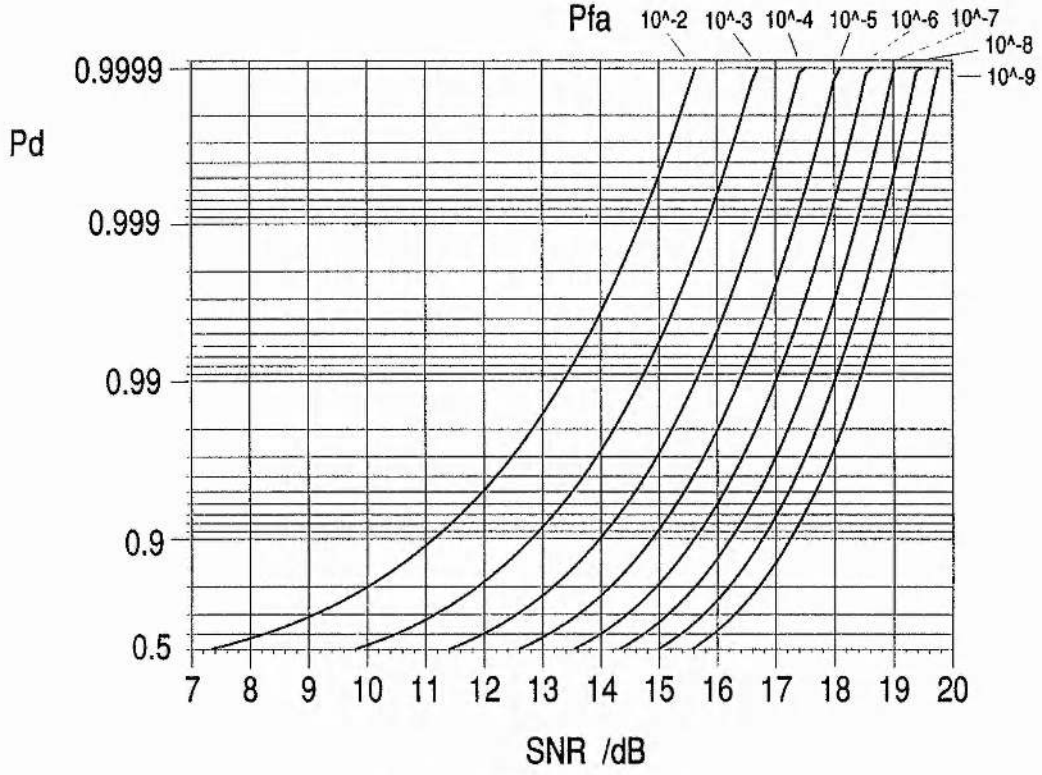


Fig. 9·14 : Probability of Detection

For the case of an N point data set, one state will give a correct correlation while the rest are like noise and each could cause a false alarm. Since N is generally large, we can take it that there are N contributions to the probability of false alarm. The probability that a false alarm occurs in a single test is P_{FA}, as shown above, so $1 - P_{FA}$ represents the single test probability of no false alarm. The probability that N tests will give no false alarm (i.e. all N tests fall below threshold) is then $(1 - P_{FA})^N$. Therefore, the probability of a false alarm in N data

points (i.e. the probability that any of the N tests is above threshold) is $1 - (1 - P_{FA})^N$.

Using this expression I calculated the false alarm probability for a $N = 4096$ point data set, as used in the experimental Beacon Detection System - see Fig. 9·15. It can be seen that in Fig. 9·15, the curve covers a narrower range of *SNR* values than the one in Fig. 9·13 and they are at a slightly higher level overall.

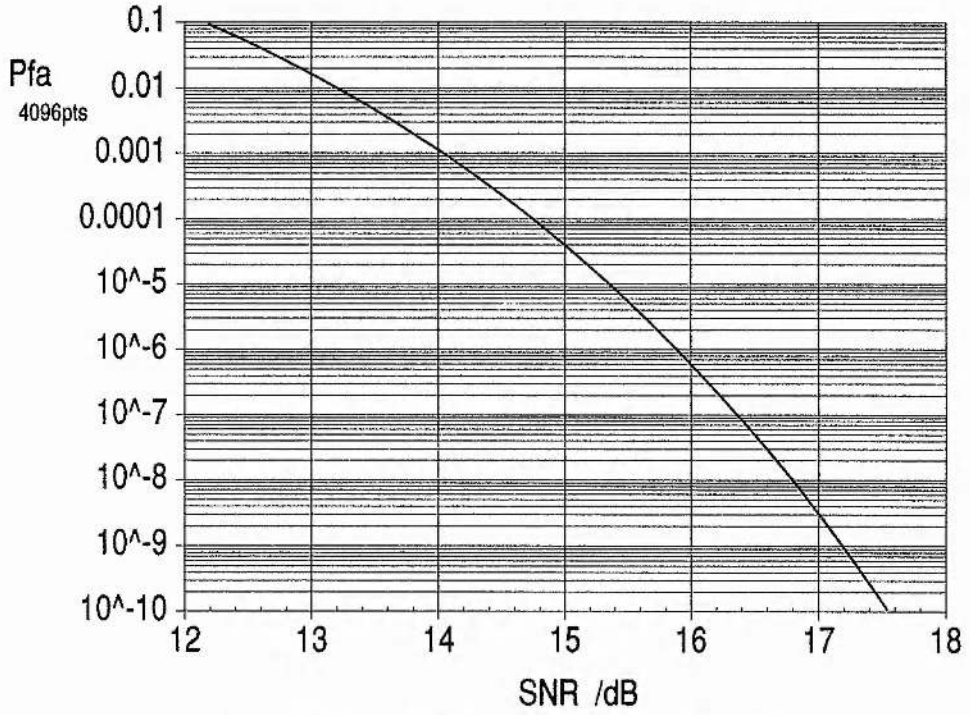


Fig. 9·15 : Probability of False Alarm for 4096 data points

Values from Fig. 9·15 were then used to generate another graph of P_D for values of P_{FA} ranging from 10^{-2} to 10^{-9}. The expression for P_D needs no amendment because in this case there is only one state which gives a correct correlation. See Fig. 9·16.

This graph is very similar to Fig. 9·14 but that the N point data set has moved the curves closer together and up to higher SNR values.

As an example, if an acceptable probability of false alarm is one in 10^7 (i.e. $P_{FA} = 10^{-7}$) and we require 99% correct identification (i.e. $P_D = 0.99$) then we need a signal to noise ratio of at least 19dB.

The rapidly varying nature of these probabilities is well demonstrated by considering the effect of slightly increasing the signal to noise ratio. Consider the case where P_D = 98% and P_{FA} = 10^{-3} - the required *SNR* is 17dB. If we increase the *SNR* by just 2dB and we keep the same P_D level, the P_{FA} decreases by five orders of magnitude to 10^{-8}. Similarly, if we keep P_{FA} fixed, we find our probability of correct detection increases (off-scale) to somewhere near 99.995%.

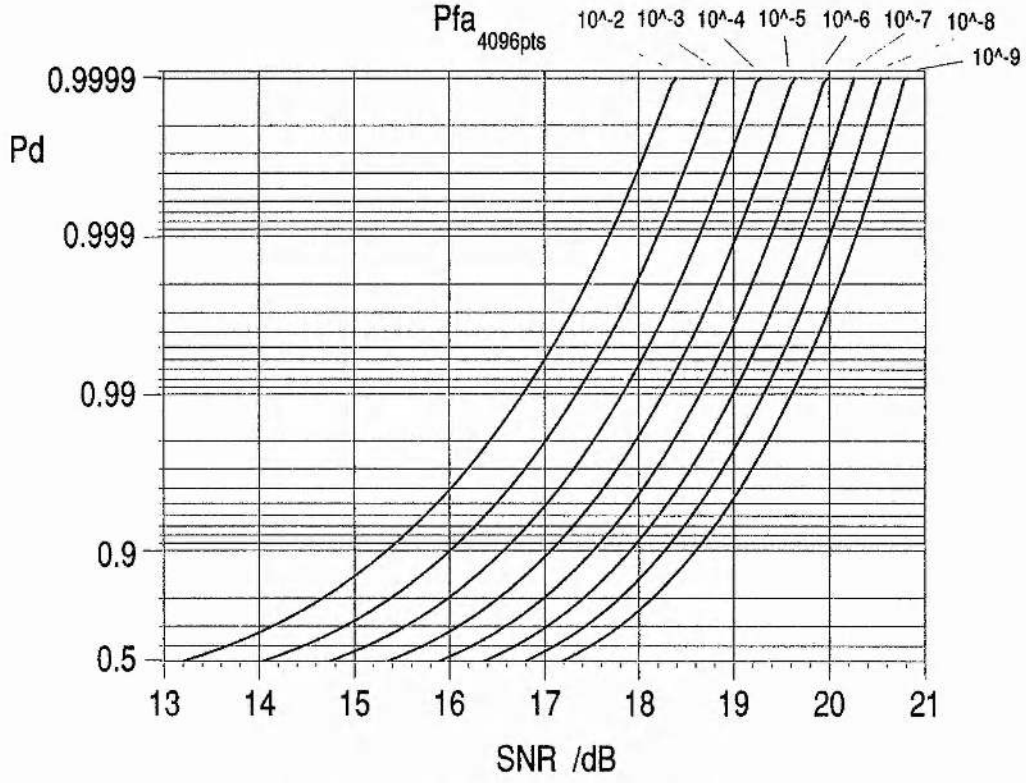


Fig. 9·16 : Probability of Detection (for 4096 noise points)

9.4 Beacon Power Requirements

Equation 9.60 derives the power level required at the receiver input ports for a given set of receiver parameters (for which the minimum output SNR is determined from the graphs in the previous section). The source power of the beacon, P_T, can then be calculated using the Link Gain equation

$$P_R = \frac{P_T L_T G_T L_R G_R}{\left(\frac{4\pi R}{\lambda}\right)^2} \qquad 9.70$$

where P_R is the received power (set by eqn. 9.60), L_R and G_R are the receiving antenna insertion loss and gain, L_T and G_T are the transmitting antenna insertion loss and gain, R is the beacon range and λ is the signal wavelength. Ideally, of course, $L_T = L_R = 1$.

Combining eqns. 9.60 and 9.70 allows one to model the required beacon power for a given receiver configuration and source range. A third version of the 'C' modelling program was used to calculate the required beacon power as a function of receiver cube size as shown in Fig. 9·17 for the case of $R = 10$ km.

System with four lenses and standard discriminator.
Underillumination d/w = 3.00 Lens diameter/cubesize = 0.900
Frequency = 94.00 GHz Discriminator arm = 20.00 mm
Tn = 3000 K Int. Time = 1.000 sec. Output S/N = 20.0dB
Bandwidth = 2.00 GHz Transmitter Gain = 5.00 dB

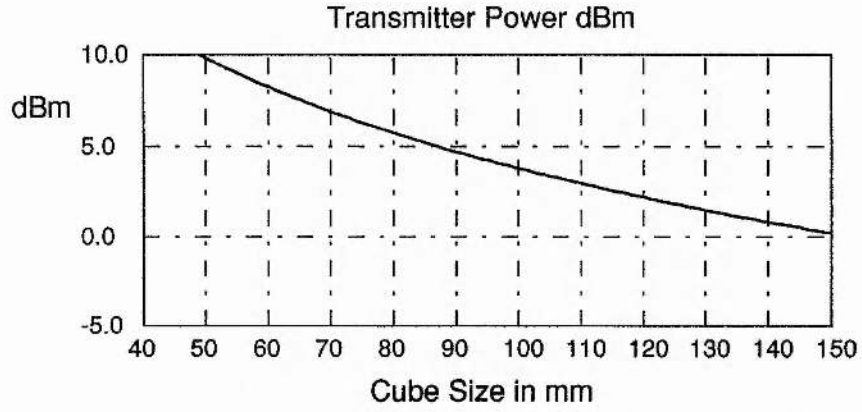


Fig. 9·17 : Required beacon power for specified receiver

One can see from Fig. 9·17 that the required power is quite small (of the order of a few milliwatts) which is easily available from Gunn oscillators. Additionally, the low power spectral density would make beacons relatively hard to detect with simple radiometric receivers and they would contribute little to the overall level of terrestrially transmitted millimetric signal power.

9.5 Real versus Optimum Receiver

Of course, a real receiver may not be able to achieve all the optimum criteria which have been assumed in the models and any shortcomings must be taken into account when assessing a working system.

The modelling of the optical performance of section 9.1 has shown that it should be possible to reduce the size of the receiver circuit substantially whilst maintaining adequate performance. Indeed, the quasi-optical circuit has been shown theoretically to offer very low levels of coupling loss and wide bandwidth operation, given a carefully optimised system. Certainly the prototype receiver, as measured, fell some way short of the possible standard but many of the shortcomings have been accounted for in the previous chapter.

Unfortunately, it has been harder to gather experimental results with which to test the predicted signal-to-noise behaviour. A verification of the statistical probabilites relating to the correct detection of correlation peaks would require many time consuming, repetitive measurements and a suitable opportunity to do this has not yet arisen.

In the future, it is my intention to continue with a detailed experimental investigation of all the aspects of receiver performance which have been considered theoretically in this chapter.

REFERENCES

[1] Wylde, R.J., "Millimetre-wave Gaussian beam-mode optics and corrugated feed horns", *Proc. IEE, Part H*, **131** (4), Aug. 1984, pp.258-262.
[2] Lesurf, J.C.G., *Millimetre-wave Optics, Devices & Systems*, Adam Hilger / IOP, 1990, p. 148.
[3] Jim Lesurf, St. Andrews University, Private Communication.
[4] Ref. 2, p. 18.
[5] Ref. 2, p. 37.
[6] Ref. 2, p. 148.
[7] Bal Kumar, Hirst Research Centre, GEC-Marconi Ltd, Borehamwood, Private Communication.
[8] Lesurf, J.C.G., *Information and Measurement*, IOP, 1995, pp. 132-133.
[9] Abramowitz, M. & Stegun, I.A., *Handbook of Mathematical Functions*, Dover Publications, 1972, eqns. 7.1.1 & 7.1.2.
[10] ibid, eqn. 7.1.25.

Chapter Ten
A Final Review

10.1 Conclusion

This thesis has set out to show that quasi-optical signal processing techniques can be applied to spread spectrum millimetre wave systems.

The properties and advantages of millimetre waves have been introduced in respect of communications applications, with emphasis on broadband techniques.

Chapters 3 and 4 focus on the use of Gunn oscillators as frequency modulated sources. In Chapter 3 a new design of Gunn oscillator power supply is presented. It has a modulation bandwidth of over 10MHz, which is about two orders of magnitude greater than the fastest pre-existing unit available at St. Andrews, and yet has a noise performance which is only a few dB worse than that of the lowest noise pre-existing power supply.

Chapter 4 looks at the bias tuning behaviour of Gunn oscillators. The results suggest that the bias tuning characteristics are primarily a function of the device rather than its embedding circuit. The bias tuning characterisation was done with a automated system running under computer control. The Auto_Bias2 system makes these measurements more reliably than when they are made manually and it also vastly reduces the time taken to complete a full oscillator characterisation.

Self-modulation effects in W-band Gunn oscillators are discussed in Chapter 5. These effects can produce nanosecond pulses of millimetric signals with relatively high peak power. The pulse effects are found to be repeatable and are a function of the bias voltage and bias line termination. Under certain conditions there is evidence for semi-chaotic effects which generate very broadband noise-like output.

The design, construction and analysis of the Beacon Detection System as recorded in Chapters 7, 8 and 9 has demonstrated the application of quasi-optical signal processing methods to a millimetre wave spread spectrum system.

Chapter 7 presents the quasi-optical frequency discriminator as applied to the demodulation and correlation of broadband signals. Experimental results confirm the theory of operation of the discriminator and correlation process. The virtues of spread spectrum process gain are demonstrated qualitatively.

In Chapter 8 I have demonstrated the technique of two-port spatial interferometry as a method of providing sub-beamwidth angular resolution. The properties of the two-port are derived mathematically and the experimental results are concurrent with theory. The angular resolution of the interferometer is of the order of one tenth of a single port beamwidth. The complete Beacon Detection System receiver also shows results which are consistent with theory. The loss in the quasi-optical circuit is measured and the various contributions are accounted for. The addition of free-space planar Faraday rotators provides an increase in optical throughput of about 2.5dB. The overall loss in the quasi-optical circuit, despite the lack of a set of correctly designed lenses, is measured as -3dB.

A theoretical analysis of the optical performance of the receiver is discussed in Chapter 9. Computer models predict that almost minimal loss is possible with a correctly optimised system. Additionally, it should be feasible to reduce the size of the quasi-optical circuit by about a factor of two without loss of performance.

Furthermore, computer simulations of the signal-to-noise behaviour of the receiver predict a very high level of receiver sensitivity. The mathematical basis for these calculations is derived in Chapter 9 along with the derivation of the probabilities of detection and false alarm for the correlator output.

Overall, the Beacon Detection System shows that quasi-optical signal processing methods are highly suitable for spread spectrum

millimetre wave applications. The prototype system demonstrates the principles successfully and that the experimental results are supported by theoretical analyses. Using computer models it should be possible to design and construct practical, very high performance millimetric systems which employ quasi-optical techniques.

10.2 Future Work

There are a number of areas of research which have been touched on in this thesis that would merit further investigation.

In order to improve Gunn oscillators as frequency modulated sources, more study of their dynamic bias tuning characteristics is required. A better understanding of the mechanisms responsible for bias tuning behaviour would allow the current imperfections to be addressed and hopefully removed.

As a related issue, I intend to investigate further the self-modulation effects in certain types of Gunn diode, as described in Chapter 5. An understanding of these mechanisms may allow one to design both devices which are free from self-modulation behaviour and those which exploit the effect to produce inherently wideband sources. The semi-chaotic effects described merit a great deal more study as the field of chaotic behaviour in millimetric components is relatively new. Experimental studies should be conducted in parallel with computer simulations of nonlinear models.

The work presented in Chapters 7, 8 and 9 suggests a great many possible future projects. It is my intention to conduct further field trials of the Beacon Detection System and continue with the experimental characterisation. This would lead to the BDS receiver being optimised. Then it should be possible to verify the theoretical levels of optical and signal-to-noise performance which were predicted by the computer models.

The quasi-optical circuit of the BDS described in this thesis is

perhaps only one way in which a spread spectrum receiver could be built. Novel quasi-optical circuits may be able to provide the required signal processing capabilities with less complexity.

Additionally, it would be desirable to test the validity of the models which project a reduction in the scale of the quasi-optical components. Hopefully, this will be achieved through the proposed construction of a spatial interferometer using cubes which are approximately half the size of those used at present. The practical realisation of more compact quasi-optical circuits would remove one of the few long-standing drawbacks to this technology. If this is possible then compact, high performance quasi-optical circuits could be utilised in a much wider variety of applications than is the case at present.

Appendix 1
Auto_Bias2 Software

The code written for controlling the automatic bias tuning measurement system described in Chapter 4 is presented here. The program runs in interpreted BASIC *V* on an Acorn 32-bit RISC computer. The listing is preceded by explanations of the procedures and functions incorporated in the code. The list is ordered alphabetically by procedure/function name.

PROCcalibrate_power allows the Anritsu monitor signal to be converted to actual power values. It permits the user to vary the preamp gain to give a reasonable voltage such that quantisation in the ADC is minimised. The user is then asked for the current power reading and a calibration factor is calculated which is subsequently used to convert the sampled monitor voltages to power readings.

PROCcool_down waits for the number of seconds specified at the start of the listing to allow the oscillator to stabilise at the start of the ramp. A visual indicator is provided in the form of white dots showing frequency against time.

PROCdac_step takes an integer, converts it to a four figure hex string and sends it to the EIP Counter as an instruction to set the DAC to the given value.

PROCdelay(n) generates an *n* second delay.

PROCeip_ready resets the EIP, asks it whether a measurement is available and waits until a frequency count has been completed.

PROCeip_start sets up the EIP counter. The Counter is fully reset and halts execution of the code until it is ready for measurements.

PROCerror1 is used in the code before data has been acquired. It resets the EIP.

PROCerror2 is used in the code during the data acquistion routines. It resets the EIP and calls the file saving procedure.

GPIB Routines are all taken from the Intelligent Interfaces library of commands and have been copied straight across.

PROCinit_values measures the nominal frequency and bias before moving to the start of the sweep.

PROCoff resets the EIP.

FNread_freq calls the EIP to read the frequency and returns the result in GHz.

FNsample_bias measures the oscillator bias voltage on ADC Ch8. The

factor 10.03 was chosen after calibrating the modified input. The result is rounded to three decimal places.

PROCsample_init sets up the Wild Vision ADC using standard SYS commands. An area of memory is set aside for the sampled data and the clock speed is set to the highest available.

FNsample_power measures the power monitor signal on ADC Ch3 and returns a value of power in milliwatts. The result is rounded to three decimal places.

PROCsave_up_file asks for a filename and saves the sweep data to disc as a text file of that name. The data is saved in the form:- *bias, frequency, power.*

PROCstartupGPIB sets up some standard parameters as recommended in the handbook of the GPIB card.

PROCsweep_to_start sweeps the oscillator bias from the nominal position to the lower end of the ramp.

PROCsweep_up increments the bias and measures frequency, bias and power at each step. A given delay is incorporated between stepping and taking measurements to allow for a certain amount of thermal stabilisation. The data are displayed on screen as it is acquired.

```
REM Auto_Bias2  v1.2
REM
REM Automated Frequency v. Power measurements
REM using EIP counter and Anritsu Meter
REM
REM © Duncan A. Robertson       27th January 1993
REM St. Andrews University
:
ON ERROR PROCerror1 :END
:
MODE 20
root$ = ":Card_Apps.$.EIP."       :REM Root directory for saved data
no_steps% = 25                    :REM Number of steps in full sweep
jump% = &3FFF/no_steps%           :REM DAC jump calculated from number of steps
ct%=0                             :REM Counting flag
stable% = 120                     :REM No. of seconds to allow osc. to stabilise
:
PROCstartupGPIB
PROCsample_init
PRINT"Automatic FREQUENCY : POWER Tuning Measurements"
PROCeip_start
:
size%=no_steps%+10
DIM FREQUP(size%)
DIM BIASUP(size%)
DIM PWRUP(size%)
:
PROCinit_values
PROCsweep_to_start
:
```

```
ON ERROR PROCerror2 :END
:
PROCsweep_up
VDU7
PROCadrandoutput(0,19,"PFP")      :REM Resets DAC to +5V
PROCsave_up_file
END
:
:

DEFPROCcool_down
PRINTTAB(0,10)"Please wait for oscillator to thermally stabilise - allow
   ";stable%;" seconds "
PROCdelay(2)
TIME = 0
lower_freq = FNread_freq
REPEAT
old_freq = FNread_freq
PRINTTAB(0,20)"Frequency = ";old_freq
yo = (old_freq-lower_freq)*8000+550
PRINTTAB(30,20)"TIME = ";TIME/100;"  seconds "        :REM Prints time elapsed
xo = TIME/15
IF xo > 1250 THEN xo-=1250
CIRCLEFILL xo,yo,2                :REM Plots dot indicating relative frequency
UNTIL TIME > stable%*100
VDU7
PROCdelay(3)
ENDPROC

DEFPROCdac_step(level%)
big% = level%+&10000              :REM
hex$ = STR$~(big%)                :REM Converts no. to a four figure hex string
dac_val$ = RIGHT$(hex$,4)         :REM
cmd$ = "TA11"+dac_val$
PROCadrandoutput(0,19,cmd$)             :REM Sets EIP DAC output
PROCadrandoutput(0,19,"TA11TEST")       :REM Clears EIP display
ENDPROC

DEFPROCdelay(n%)                        :REM Delay in seconds
TIME=0:REPEAT:UNTIL TIME>n%*100
ENDPROC

DEFPROCeip_ready
PROCadrandoutput(0,19,"RS")             :REM Resets EIP
PROCadrandoutput(0,19,"SR01")           :REM Measurement available?
REPEAT
state% = FNservice_requested            :REM Waits until count completed.
UNTIL state% = TRUE
ENDPROC

DEFPROCeip_start
PROCadrandoutput(0,19,"PFP")            :REM Resets DAC to +5V
PROCeip_ready
PRINT"EIP ready "
PROCdelay(2)
PROCadrandoutput(0,19,"ES")             :REM Scientific Notation
ENDPROC

DEFPROCinit_values                      :REM Records some starting values
PROCeip_ready
start_freq = FNread_freq
nominal_bias = FNsample_bias
```

```
ENDPROC

DEFPROCoff
PROCadrandoutput(0,19,"RS")              :REM Reset EIP
ENDPROC

DEFFNread_freq
PROCadrandoutput(0,19,"HA")              :REM  Hold Active
PROCadrandoutput(0,19,"FR")              :REM  MEASURES FREQ
result$ = FNadrandinput(0,19,255)        :REM  Reads Frequency
freq = VAL(result$)/1000000000           :REM  Converts to Gigahertz
PROCadrandoutput(0,19,"HP")              :REM  Hold Passive
= freq

DEFPROCsave_up_file
PRINT
PRINT"UPWARD Sweep Data Saving Routine"
PRINT
INPUT"Please input filename ";file$
address$ = root$+file$
channel = OPENOUT(address$)              :REM Creates data file on disk
PRINT"Saving to disc..."
FOR loop% = 0 TO ct%-1
BPUT# channel, STR$(BIASUP(loop%))+","+STR$(FREQUP(loop%))+","+STR$(PWRUP(loop%))
NEXT
CLOSE# channel
OSCLI("SETTYPE "+address$+" TEXT")
PRINT"                    .........Finished!"
ENDPROC

DEFPROCstartupGPIB
PROCeos_out_1chr(13)                :REM END OF STRING OUTPUT CHARACTER
PROCeos_in_1chr(13)                 :REM END OF STRING INPUT CHARACTER
PROCtimeout_on                      :REM ENABLES TIMEOUTS
PROCinit                            :REM INITIALISES HARDWARE
ENDPROC

DEFPROCsweep_to_start
PRINT"Moving to start of ramp"
FOR stage% = &1FFF TO &3FFF STEP &20
PROCdac_step(stage%)                      :REM Set DAC to XXXX volts
NEXT
PROCeip_ready
PROCdelay(3)
PROCcool_down
ENDPROC

DEFPROCsweep_up
CLS
PRINT"Commencing Sweep Measurements"
PROCeip_ready
PRINT"EIP ready to sweep"
PROCdelay(2)
ct%=0
PRINTTAB(0,20)"P"                   :REM Labels screen axes
PRINTTAB(0,21)"o"
PRINTTAB(0,22)"w"
PRINTTAB(0,23)"e"
PRINTTAB(0,24)"r"
PRINTTAB(0,45)"F"
PRINTTAB(0,46)"r"
PRINTTAB(0,47)"e"
```

```
PRINTTAB(0,48)"q"
PRINTTAB(0,49)"u"
PRINTTAB(0,50)"e"
PRINTTAB(0,51)"n"
PRINTTAB(0,52)"c"
PRINTTAB(0,53)"y"
PRINTTAB(20,59)"Bias "
FOR ramp% = &3FFF TO &0000 STEP -jump%
PROCdac_step(ramp%)                                    :REM Increment DAC
PROCeip_ready                                          :REM Waits for EIP to count
PROCdelay(15)                                          :REM Waits for 15 seconds
freq = FNread_freq                                     :REM Reads frequency
bias_volts = FNsample_bias                             :REM Reads bias
power_monitor = FNsample_power                         :REM Reads power
PRINTTAB(0,5)"Frequency = ";freq;"  GHz  "
PRINTTAB(20,5)"  Bias = ";bias_volts;" V  "
PRINTTAB(50,5)"Power = ";power_monitor;" mW  "
x_bias = ((bias_volts)-nominal_bias)*550+400
y_power =  power_monitor*50
y_freq = (freq-old_freq)*1000+100
CIRCLEFILL x_bias,y_power,2              :REM Displays data as it is acquired
CIRCLEFILL x_bias,y_freq,2
FREQUP(ct%)=freq                         :REM
BIASUP(ct%)=bias_volts                   :REM Stores data in arrays
PWRUP(ct%)=power_monitor                 :REM
ct%+=1
statusbyte% = FNserial_poll(19)
NEXT
VDU7
ENDPROC

DEFPROCerror1
PROCoff
PROCadrandoutput(0,19,"DN")              :REM RESETS DISPLAY TO NORMAL
PROCadrandoutput(0,19,"PFP")             :REM Resets Coarse Tune to +5V
REPORT:PRINT" at line ";ERL
ENDPROC

DEFPROCerror2
PROCoff
PROCadrandoutput(0,19,"DN")              :REM RESETS DISPLAY TO NORMAL
PROCadrandoutput(0,19,"PFP")             :REM Resets Coarse Tune to +5V
PRINT
PRINT
PRINT
PROCsave_up_file
REPORT:PRINT" at line ";ERL
ENDPROC
:

REM ***************************************************
REM ************  GPIB ROUTINES BELOW  ****************
REM ***************************************************

DEF PROCeos_out_1chr(eos1%)
SYS &40340,&00,&03,&01,eos1%
ENDPROC

DEF PROCeos_in_1chr(eos1%)
SYS &40340,&00,&04,&01,eos1%
ENDPROC
```

```
DEF PROCtimeout_on
SYS &40340,&00,&06
ENDPROC

DEF PROCinit
LOCAL d0%,d1%,ma%
SYS &40340,&00,&02 TO d0%,d1%,ma%
IF (ma% AND &80) THEN SYS &40340,&00,&11,ma% ELSE SYS &40340,&00,&10,ma%
ENDPROC

DEF PROCadrandoutput(mode%,lad%,data$)
SYS &40340,&00,&08,mode%
SYS &40340,&00,&3C,&00,lad%,&00,data$,LEN(data$)
ENDPROC

DEF FNadrandinput(mode%,tad%,bytecnt%)
LOCALd0%,d1%,d2%,d3%,d4%,databuf$,d6%,xcnt%
IFbytecnt%>254THEN bytecnt% = 254
databuf$ = STRING$(bytecnt%," ")
SYS &40340,&00,&09,mode%
SYS&40340,&00,&3D,&00,tad%,
   &00,databuf$,bytecnt%TOd0%,d1%,d2%,d3%,d4%,databuf$,d6%,xcnt%
=LEFT$(databuf$,xcnt%)

DEF FNserial_poll(tad%)
LOCAL d0%,d1%,stb%
SYS &40340,&00,&2B,&00,tad% TO d0%,d1%,stb%
=stb%

DEF FNservice_requested
LOCAL d0%,d1%,d2%,d3%,bussts%,status%
SYS &40340,&00,&00 TO d0%,d1%,d2%,d3%,bussts%
IF (bussts% AND &04) = 0 THEN status%=FALSE ELSE status% = TRUE
=status%

REM  ************************************************
REM  ********  Wild Vision Routines Below  ***********
REM  ************************************************

DEFPROCsample_init
Limit%=1024                :REM  Sets the data set size
DIM Buffer% Limit%*4       :REM  Assigns memory area called Buffer% - 4 times
                           :REM  bigger because words
:
REM=======================================================
REM By converting the SWI names into numbers we save time
REM in critical routines later.
REM=======================================================
:
SYS "OS_SWINumberFromString",,"WVAIO_Clock"      TO Clock%
SYS "OS_SWINumberFromString",,"WVAIO_Channel"    TO Channel%
SYS "OS_SWINumberFromString",,"WVAIO_ADC"        TO ADC%
:
SYS Clock%,160000                :REM  Maximum sampling rate (167kHz)
PROCcalibrate_power
ENDPROC

DEFPROCcalibrate_power
SYS Channel%,3                   :REM  Set channel number
SYS ADC%,1,1,Buffer%             :REM  Makes 1 acquisition from 1 channel

PRINT"**********    POWER CALIBRATION      ***********"
```

```
PRINT
PRINT
PRINT"Adjust preamp gain until signal is about 2.5 - 3 volts"
PRINT
PRINT"Then press ANY key to continue"
REPEAT
SYS ADC%,1,1,Buffer%                    :REM  Makes 1 acquisition from 1 channel
value=((!Buffer%)-2048)/2048)*5         :REM  Calculates voltage from ADC value
int%=value*100                          :REM
ans=int%/100                            :REM  Rounds off voltage to 2 dec. pl.
PRINTTAB(0,10)"Voltage of signal =   ";ans;"   Volts     "
A$ = INKEY$(0)
UNTIL A$ <> ""
PRINT
INPUT"What is the current Power Reading ";reading
units$="mW"
cal_factor = (reading/value)
PRINTTAB(45,10)"Calibration Factor = ";cal_factor
power = value*cal_factor                                :REM  Calibrates power
PRINTTAB(45,12)"Power Reading = ";power;" ";units$
PROCdelay(4):VDU7:CLS
ENDPROC

DEFFNsample_bias
SYS Channel%,8                          :REM  Set channel number
SYS ADC%,1,1,Buffer%                    :REM  Makes 1 acquisition from 1 channel
vol = (((!Buffer%)-2048)/2048)*10.03    :REM  Calculates voltage (scaling=10.03)
vol% = vol*1000                         :REM
voltage = vol%/1000                     :REM  Rounds off bias to 3 dec. pl.
=voltage

DEFFNsample_power
SYS Channel%,3                          :REM  Set channel number
SYS ADC%,1,1,Buffer%                    :REM  Makes 1 acquisition from 1 channel
pow = (((!Buffer%)-2048)/2048)*5*cal_factor
pow% = pow*1000
power = pow%/1000                       :REM  Rounds off power to 3 dec. pl.
=power
```

Appendix 2
Correlation Software

The code written for calculating correlations as described in Chapter 7 is presented here. The program runs in interpreted BASIC *V* on an Acorn 32-bit RISC computer. The listing is preceded by explanations of the procedures and functions incorporated in the code. The list is ordered alphabetically by procedure/function name.

PROCbit_rev is required for the FFT routine. It is translated direct from *Numerical Recipes in FORTRAN*, by Press, Flannery, Teukolsky & Vetterling, Cambridge Press.

PROCcomplex_mult performs the complex multiplication of the two Fourier transformed data sets.

PROCfft() calculates the FFT of the data set. If the argument is 1 the FFT is forward, whereas it is an inverse FFT if the argument is -1. It is translated directly from *Numerical Recipes in FORTRAN*. The data are stored in an array X( ) in sequential real and imaginary pairs i.e. real1, imag1, real2, imag2, ...

PROCget_data loads the data file to be correlated.

FNinterpolate interpolates a new data point from a specified range of existing points.

PROCload_dataset_1 loads the data set from real and imaginary arrays into the FFT array X( ).

PROCload_FFT loads the stored FFT data from real and imaginary arrays into the FFT array X( ).

PROCload_product loads the result of the complex multiplication from real and imaginary arrays into the FFT array X( ).

PROCprepare_REF takes the input data, interpolates new values, removes any dc offset and scales the data by dividing by the rms value.

PROCprocess is the main procedure which calls the various stages of the correlation calculation.

PROCsave_correlation saves the correlation data to a text file.

PROCsaveFFT_data saves the FFT of a reference data set for later use by the correlation routine.

FNsinc() calculates the sinc of a value

PROCsinc_interp() handles the sinc interpolation routine for a given size of data set.

PROCsnr calculates the output signal to nose ratio of the correlation.

PROCstartup defines all the necessary arrays and variables.

PROCstore_first_FFT stores the result of the FFT of the dataset into real and imaginary arrays.

PROCstore_inverse_FFT stores the result of the inverse FFT into real and imaginary arrays for extracting the correlation data.

```
REM Correlation Software
REM
REM © Duncan A. Robertson
REM St. Andrews University
:
PROCstartup
PROCload_FFT
PROCprocess
END
:
:
DEFPROCprocess
SYS "Hourglass_On"
PROCget_data
PROCprepare_REF
PROCload_dataset_1
PROCfft(1)                          :REM  This calls for a forward FFT
PROCstore_first_FFT
 REM PROCsaveFFT_data               :REM These two lines are used when saving a
 REM ENDPROC                        :REM  reference FFT
PROCcomplex_mult
PROCload_product
PROCfft(-1)                         :REM  This calls for an inverse FFT
PROCstore_inverse_FFT
PROCsnr
PROCsave_correlation
SYS "Hourglass_Off"
ENDPROC
:
DEFPROCcomplex_mult
FOR z% = 1 TO nn%
PRR(z%) = (FR2(z%)*FR(z%))+(FI2(z%)*FI(z%))
PRI(z%) = (FR(z%)*FI2(z%))-(FR2(z%)*FI(z%))
NEXT
ENDPROC
:
DEFPROCstore_first_FFT
z% = 1
FOR q% = 1 TO n% STEP 2
FR(z%) = X(q%)
FI(z%) = X(q%+1)
z% += 1
NEXT
ENDPROC
```

```
:
DEFPROCstore_inverse_FFT
z% = 1
FOR q% = 1 TO n% STEP 2
FR(z%) = X(q%)/nn%
FI(z% )= X(q%+1)/nn%
z% += 1
NEXT
ENDPROC
:
DEFPROCfft(isign%)
PROCbit_rev
mmax% = 2
WHILE n% > mmax%
istep% = 2*mmax%
theta = 2*PI/(isign%*mmax%)
wpr = -2*SIN(0.5*theta)^2
wpi = SIN(theta)
wr = 1
wi = 0
FOR m% = 1 TO mmax% STEP 2
FOR i% = m% TO n% STEP istep%
j% = i% + mmax%
tr = wr*X(j%)-wi*X(j%+1)
ti = wr*X(j%+1)+wi*X(j%)
X(j%) = X(i%)-tr
X(j%+1) = X(i%+1)-ti
X(i%) = X(i%)+tr
X(i%+1) = X(i%+1)+ti
NEXT
wtemp = wr
wr = wr*wpr-wi*wpi+wr
wi = wi*wpr+wtemp*wpi+wi
NEXT
mmax% = istep%
ENDWHILE
IF isign% = -1 THEN
X() = X()/nn%
ENDIF
ENDPROC
:
DEFPROCprepare_REF
PROCsinc_interp(REF())
REF() = STRETCH()                    :REM Interpolates the data set
ave = 0
max = -1E6
min = 1E6
FOR inow% = 1 TO nn%
now = REF(inow%)
ave += now
IF now > max : max = now
IF now < min : min =now
```

```
NEXT
ave = ave/nn%
REF() = REF()-ave                        :REM Suppresses dc
rms = 0
FOR inow% = 1 TO nn%
now = REF(inow%)
rms += now*now
REF(inow%) = now
NEXT
rms = SQR(rms/nn%)
REF() = REF()/rms                        :REM Scales data set
ENDPROC
:
DEFPROCsinc_interp(SAMPLED())
dto% = 10                          :REM  old data interval in microseconds
dtor = dto%                        :REM  real value of dto%
dtn% = 8                           :REM  new data interval in us
K% = 2                             :REM  Interpolation range (multiple of 2)
FOR loop% = 1 TO nn%
t% = loop%*dtn%
tr = t%                            :REM  real value of time
STRETCH(loop%) = FNinterpolate
NEXT
ENDPROC
:
DEFFNinterpolate
sigma = 0
td_midpt% = (t% DIV dto%) + (K%/2)
 FOR ti_pos% = td_midpt%-(K%/2) TO td_midpt%+(K%/2)
  ti = ti_pos%*dtor
  exprn = (PI*(tr-ti))/dtor
  factor = FNsinc(exprn)
  sigma += factor*SAMPLED(ti_pos%)
 NEXT
=sigma
:
DEFFNsinc(zz)
LOCAL zz,result
IF ABS(zz) < 1E-5 THEN
  result = 1
ELSE result = (SIN(zz))/zz
ENDIF
=result
:
DEFPROCbit_rev
j% = 1
FOR i% =1 TO n% STEP 2
IF j% > i% THEN
tR = X(j%)
tI = X(j%+1)
X(j%) = X(i%)
X(j%+1) = X(i%+1)
```

```
X(i%) = tR
X(i%+1) = tI
ENDIF
m% = n%/2
WHILE m% >= 2 AND j% > m%
j% = j%-m%
m% = m%/2
ENDWHILE
j% = j%+m%
NEXT
ENDPROC
:
DEFPROCsnr
cmax = -1000
crms = 0
FOR q% = 1 TO nn%
IF FR(q%) > cmax THEN
    cmax = FR(q%)
    index% = q%
    ENDIF
NEXT
cmax = FR(index%)
FOR inow% = 1 TO (index%-10)
cnow = FR(inow%)
crms += cnow*cnow
NEXT
FOR inow% = (index%+10) TO nn%
cnow = FR(inow%)
crms += cnow*cnow
NEXT
crms = SQR(crms/(nn%-20))
snr = FR(index%)/crms
dB = 10*LOG(snr)
ENDPROC
:
DEFPROCsave_correlation
address$ = save_name$
channel = OPENOUT(address$)
BPUT# channel,"Title, max = "+STR$(cmax)+" at T = "+STR$(index%)+"  rms = "+
STR$(crms)+" SNR = "+STR$(snr)+" ("+STR$(dB)+"dB)"
FOR q% = 1 TO nn%
BPUT# channel, STR$(q%)+","+STR$(FR(q%))
NEXT
CLOSE# channel
OSCLI("SETTYPE "+address$+" TEXT")
ENDPROC
:
DEFPROCsaveFFT_data
address$ = "FFTdata"
channel = OPENOUT(address$)
FOR q% = 1 TO nn%
BPUT# channel, STR$(FR(q%))+","+STR$(FI(q%))
```

```
NEXT
CLOSE# channel
OSCLI("SETTYPE "+address$+" TEXT")
ENDPROC
:
DEFPROCget_data
j% = 1
file% = OPENIN pathname$
REPEAT
this$ = GET$# file%
REF(j%) = VAL(this$)
j%+ = 1
UNTIL j% > nn%
CLOSE# file%
ENDPROC
:
DEFPROCload_dataset_1
z% = 1
FOR q% = 1 TO nn%
X(z%) = REF(q%)
X(z%+1) = 0
z% += 2
NEXT
ENDPROC
:
DEFPROCload_product
z% = 1
FOR q% = 1 TO nn%
X(z%) = PRR(q%)
X(z%+1) = PRI(q%)
z% += 2
NEXT
ENDPROC
:
DEFPROCstartup
nn% = 4096                      :REM Number of points in data set
DIM REF(nn%)                    :REM Reference data set
DIM SIG(nn%)                    :REM Signal data set
DIM SAMPLED(nn%)                :REM Original data set before interpolation
DIM STRETCH(nn%)                :REM Data set after interpolation
DIM FR(nn%)                     :REM Real part of data set
DIM FI(nn%)                     :REM Imaginary part of data set
DIM FR2(nn%)                    :REM Real part of reference set
DIM FI2(nn%)                    :REM Imaginary part of reference set
DIM PRR(nn%)                    :REM Real part of complex product
DIM PRI(nn%)                    :REM Imaginary part of complex product
n% = 2*nn%                      :REM Size of FFT array X() - twice data set size
DIM X(n%)                       :REM FFT array
ENDPROC
:
```

```
DEFPROCload_FFT
SYS "Hourglass_On"
j% = 1
file% = OPENIN "FFTdata"
REPEAT
this$ = GET$# file%
co% = INSTR(this$,",",0)
fi$ = LEFT$(this$,co%-1)
se$ = RIGHT$(this$,(LEN(this$)-co%))
FR2(j%) = VAL(fi$)
FI2(j%) = VAL(se$)
j% += 1
UNTIL j% > nn%
CLOSE# file%
SYS "Hourglass_Off"
ENDPROC
```

::

Appendix 3
Probability Software

The programs written for calculating the probabilities of false alarm and detection referred to in Chapter 9 are presented here. They run in interpreted BASIC *V* on an Acorn 32-bit RISC computer. The theory behind the code is explained in Chapter 9.

```
REM  Probability of False Alarm in Gausssian Noise
REM
ON ERROR REPORT:PRINT " At Line ";ERL:END
PROCinit
PROCcalc
END
:
:
DEFPROCcalc
address$ = "RAM::RamDisc0.$.AnsFA_N"
channel = OPENOUT(address$)
FOR SNR=start TO stop STEP 0.005
arg=2*(SNR^2)
value=1-(1-0.5*(1-(FNerf(SNR))))^N
BPUT# channel, STR$(10*LOG(arg))+" , "+STR$((value))
NEXT
CLOSE# channel
OSCLI("SETTYPE "+address$+" &dfe")
ENDPROC
:
DEFPROCinit
start=2.877
stop=5.376
N=4096
P=0.47047
a1=0.3480242
a2=-0.0958798
a3=0.7478556
ENDPROC
:
DEFFNerf(z)
t=1/(1+P*z)
poly=a1*t+a2*t^2+a3*t^3
=1-poly*EXP(-1*z^2)
::
```

```
REM  Probability of Detection in Gausssian Noise
REM  for given levels of Probability of False Alarm
REM  ALL CALCULATIONS IN VOLTS!!!!!!!
REM
ON ERROR REPORT:PRINT " At Line ";ERL:END
PROCinit
PROCcalc
END
:
:
DEFPROCcalc
address$ = "RAM::RamDisc0.$.AnsWork2N"
channel = OPENOUT(address$)
SNRdB=start
REPEAT
SNRreal = 10^(SNRdB/10)
SNRvolts=SQR(SNRreal)
value=FNProbDet((Tvolts-SNRvolts)/1.414)
BPUT# channel, STR$(SNRdB)+" , "+STR$(value)
SNRdB+=0.05
UNTIL value>0.9999
CLOSE# channel
OSCLI("SETTYPE "+address$+" &dfe")
ENDPROC
:
DEFPROCinit
SNRreal=0
SNRdB=0
N=4096
P=0.47047
a1=0.3480242
a2=-0.0958798
a3=0.7478556
INPUT"What threshold level do you want, in dB ";TdB
Treal = 10^(TdB/10)
Tvolts=SQR(Treal)
Pfa=1-(1-0.5*(1-FNerf(Tvolts/1.414)))^N
start=TdB
stop=20
ENDPROC
:
DEFFNProbDet(w)
=0.5*(1-FNerf(w))
:
DEFFNerf(q)
z=ABS(q)
t=1/(1+P*z)
poly=a1*t+a2*t^2+a3*t^3
erf=1-poly*EXP(-i*z^2)
=SGN(q)*erf
::
```